DOCUMENT RESUME

ED 104 933 TM 004 389

AUTHOR Hall, Charles E.; And Others
TITLE VARAN2: A Linear Model Variance Analysis Program. Second Edition. Research Memorandum.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO RM-73-30
PUB DATE Nov 73
NOTE 124p.; Not available in hard copy due to marginal legibility of originality document

EDRS PRICE MF-$0.76 HC Not Available from EDRS..PLUS POSTAGE
DESCRIPTORS *Analysis of Variance; *Computer Programs; *Correlation; Discriminant Analysis; Electronic Data Processing; Factor Analysis; Hypothesis Testing; Input Output; *Manuals; Mathematical Applications; Matrices; Orthogonal Rotation; Programing Languages; *Statistical Analysis; Tests of Significance

ABSTRACT

The VARAN (variance Analysis) program is an addition to a series of computer programs for multivariate analysis of variance. The development of VARAN exploits the full linear model. Analysis of variance, univariate and multivariate, is the program's main target. Correlation analysis of all types is available with printout in the vernacular of correlation. The hybrid of these, homogeneity of regression, has been added with as much flexibility as can be currently mustered. In addition to these, VARAN includes several styles of factor and component analysis complete with tests of factorization and rotation techniques. This research memorandum is the manual for the second edition of VARAN, an enlargement of the first edition. The mainstream of the program is essentially unchanged but several additions have been made and three small programming errors have been corrected. The most extensive addition has been in serial correlation analysis. (Author/RC)

RESEARCH MEMORANDUM

RM-73-30

VARAN2:

A LINEAR MODEL VARIANCE ANALYSIS PROGRAM

Second Edition

Charles E. Hall

Katherine Kornhauser

and

Dorothy T. Thayer

This Memorandum is for interoffice use. It is not to be cited as a published report without the specific permission of the authors.

Educational Testing Service
Princeton, New Jersey
November 1973

Index

Introduction

This memorandum is the manual for the VARAN program. (The acronym comes from the words VARiance ANalysis)

The VARAN program is the latest addition to a series of computer programs for multivariate analysis of variance which originated about 1957. About that time R. Darrell Bock had a MANOVA type program running on the UNIVAC at the University of North Carolina. Bock's original program handled only a few variables and was otherwise quite restricted and was not in much general use. Later, in 1962, C. E. Hall and Elliot Cramer with Bock's assistance published a program called MANOVA in FORTRAN 2 which was internationally circulated. The program handled 25 variables and 100 degrees of freedom for hypothesis.

The wide use of this program prompted further development by Cramer and by Bock and Finn. In 1966 Cramer published a program called MANOVA in FORTRAN 4 followed by Finn and Bock's program MULTIVARIANCE also in FORTRAN 4. The development of these two programs greatly increased the scope of multivariate analyses which could be performed on computers.

MULTIVARIANCE was the first of the programs in this series that utilized the full linear model. The earlier programs had been restricted to models of the cell means in their main streams of calculation. Other variations on the linear model, like canonical correlation, were of an "accidental" nature. With the development of MULTIVARIANCE, the series turned to the flexibility of the full linear model.

The development of VARAN continues this series in the exploitation of the full linear model. Analysis of variance, univariate and multivariate, is the main target of the program as with the earlier programs. Correlation analysis of all types is available with printout in the vernacular of correlation. The hybrid of these, homogeneity of regression, has been added with as much flexibility as can be currently mustered. In addition to these, VARAN includes several styles of factor and component analysis complete with tests of factorization and rotation techniques.

The addition of factor analysis and correlation techniques to a MANOVA program brings several new capacities to multivariate analysis in addition to the customary univariate and multivariate ANOVA, correlation and homogeneity of regression problems. Some of the new capacities for analyses are listed below.

1. Three kinds of battery reduction procedures from multiple correlation applied to discriminant analysis, multivariate analysis of variance and canonical correlation.

2. Seven kinds of rotation schemes from factor analysis applied to discriminant analysis, multivariate analysis of variance and canonical correlation.

3. Factor extension from factor analysis applied to discriminant analysis, multivariate analysis of variance and canonical correlation.

4. Estimation of population variance and covariance parameters from several samples as found in analysis of variance applied to correlation and factor analysis models.

5. Use of "dummy parameters" from analysis of variance applied to factor analysis, analysis of variance (obtaining intraclass correlation coefficients) and correlation analysis (complex biserial correlation).

6. Homogeneity of regression techniques from analysis of covariance applied to correlation analysis from multiple samples and factor analysis from multiple examples.

7. Interaction tables for ANOVA models.

8. Variance reduction analysis for determining the effects of non-orthogonality in ANOVA analysis.

9. Dimension reduction analysis for determining overlap of effects in MANOVA and complex canonical correlation analyses.

The VARAN program was written more in the style of a mathematical exercise in linear calculations than as a solution for specific statistical models. Therefore, the user can expect to find applications in analysis which were not specifically contemplated by the authors. The VARAN program was also constructed to be easily expanded to include new linear model techniques as they are generated. Updates will be forthcoming periodically and for this reason it is suggested that all copies be obtained either from Educational Testing Service or from the authors.

As is customary with programs of this size, the author makes no ironclad claims of arithmetic accuracy. Any errors discovered by the users will be quickly corrected and distributed to other users. It should be noted that there were 85 or more problems run to check the accuracy of the main streams of calculation. It may also be noticed that the ability to make linear transformations of the data provides a wide variety of internal checks on calculation accuracy.

All arithmetic is single precision except for orthogonal polynomial construction. The user is warned that lengthy manipulations of highly correlated variables are not advised. (This is not much of a deficiency in handling ANOVA designs since data are not generally useful when highly correlated and cell counts are generally close to orthogonality.)

The authors wish to express their appreciation to the programming staff of Educational Testing Service for three years of assistance.

Introduction to the Second Edition

The second edition of VARAN is an enlargement of the first edition. The main stream of the program is essentially unchanged but several additions have been made and three small programming errors have been corrected. (No errors have been found in the arithmetic.)

The most extensive addition has been in serial correlation analysis. This addition should prove useful in sociological surveys and in complex sampling designs.

A feature has been added which makes Potthoff and Roy models easier to use.

An orthogonal Procrustes rotation scheme has been added to aid specific hypothesis testing in canonical analysis.

Descriptions of three machine-dependent subroutines have been added to the manual to ease conversion to other machines.

The test problems have been reworked to provide tutelage as well as checks at new installations.

The manual has been added to and revised only slightly for VARAN2. None of the program set up has been deleted although additions have been made. The old manual still works for the new program.

It is hoped by the authors that these additions will make VARAN2 even more useful.

Control Cards in Brief

1. TITLE Cards: (optional)

Cols	
1-4	TITL
5-80	Any alphameric description As many cards may be used as desired

2. PROBLEM Card: (required)

Cols	
1-4	PROB
5-6	Number of variable format cards for data (10 or fewer)
8	Number of contrast card sets (at least 2, at most 7)
10	Number of individual significance test control cards (at most 6)
12	Print cell means and variances (1 means yes)*
14	Print reduced model matrix (1 means yes)
17-18	Data file number if not cards
20	Print only estimates or raw regression coefficients (1 means yes)
22	Controls printed output (Ø gives minimal output). See Table I
24	Controls printed output (Ø adds nothing to above). See Table I
26	Type of covariance adjustment (Ø or 1 does the classical adjustment; 2 uses "error" regression weights for reduction)
28	Do not copy data input file for reanalysis (1 means don't)

3. Contrast Card Set: (At least two required)

Cols	
1	Letter identification of the factor (W is not allowed); V is allowed only for continuous variables

*blank always means "no"

Cols	
2-3	Contrast code for design variables = Ø Regular contrasts (also called nominal or deviation) = 1 Special contrasts = 2 Orthogonal polynomials (integral values) (19 levels or less) = 3 Reverse Helmert contrasts = 4 Special design parameters = 5 Orthogonal polynomials (normalized values) (39 levels or less) Function code for continuous variables (after V) = Ø do nothing = N obtain Nth power of all variables (N > 1)
4-6	Number of levels of a factor (at most 40) or, after V, the number of continuous variables (at most 39)
8	Read in variable names (1 means yes)
10	Controls linear transformation of the variables = 1 Linear transformations of the variables to be read in = 2 Weighting matrix for transformations followed by linear transformations (Potthoff and Roy models) = 3 Inverse of weighting matrix for transformations (Potthoff and Roy models)
2 col fields	Number of variables in each partition (if no partitioning, do nothing) (at most NORD partitions)
Next 2 col field	,, (two commas required)
Next 2 col field	Recoding of all the factor identification or, after V, reordering of all the variables (if no recoding, or reordering, do nothing)
Next 2 col field	,, (two commas required)
20 col field	Any description of the contrast or variables
1 col	. (a period required)

4. Contrast Card Set Options: (optional)

   a. Variable names cards. FORMAT (10A8)

      Any names in 8 column fields, 10 names per card. The name cards must precede other options.

   b. Special contrasts matrices or special design parameters. FORMAT (8F10.0)

      Contrast or design matrices are entered a row at a time, the row for the grand mean being first. Contrast and design matrices are square and of dimension equal to the number of levels of the factor.

   c. Orthogonal polynomial metric. FORMAT (8F10.0)

      The coefficients of the linear polynomial fit only.

   d. Linear transformations and weighting matrices. FORMAT (8F10.0)

      Weighting matrices for transformations (col 10 = 2 or 3) precede the transformations. Each row of a weighting matrix is punched in 10 column fields, 8 entries per card on as many cards as necessary. Each row must begin on a new card. The weighting matrix must be nonsingular, square and of the order of the number of variables.

      For linear transformations (col 10 = 1, 2 or 3), each regression coefficient is punched in a 10-column field, 8 coefficients per card, as many cards as necessary for each regression equation. Each equation begins on a new card. The transformation matrix must be nonsingular square and of order equal to the total number of continuous variables.

5. Significance Test Card (required)

This card contains the model statement and indicates the tests to be calculated. The model statement is limited to at most 9 cards. Some of the more common models are these.

   a. Two way factorial analysis of variance
      ANOVA:W=Ø,A=V,B=V,AB=V.

   b. Correlation analysis
      CORREL:W=Ø,V1=V2.

   c. One way analysis of covariance
      ANOVA:W=Ø,A=V1/V2.

d. Minres factor analysis with varimax rotation
FACTOR8#2:W=Ø,Ø=V.

e. Interaction Tables in a three way factorial
INTABL:WAB=V,WAC=V,WBC=V.

f. Serial correlations between two categories, B, pooled across two samples, A.
SERCOR:W=0,A=0;BWA=V.

The following symbols are used:

ANOVA, CORREL, FACTOR, SERCOR, INTABL and VARED are acronyms to begin the card and state the kind of model

: to separate acronym from model statement

Letters are used to denote ANOVA-type "main effects"

Letter followed by a numeral for partitions (except after W)

A sequence of letters or letters and numerals without symbols between for interaction effects

+ to indicate pooling of effects

= to separate hypothesis variables from error variables.

In this description any variable or variables to the left of an equal sign is called a "hypothesis" variable; any variable or set of variables to the right of the equal sign is called an "error" variable except if it is also to the right of a slash (/) or an ampersand (&).

- (minus) for dimension reduction

, (comma) to separate tests

; (semicolon) in serial correlation: Any hypothesis variables to the left are to be treated as contrasts among sample means, any hypothesis variables to the right are to be treated as contrasts among categories

/ to indicate that covariates follow

& to indicate that extension variables follow

* in an error term between two sets of variables to indicate that the regression of the left set on the right set is to form an error term for an analysis

Ø (a zero) as an error term indicates that the hypothesis term is to be included in the model but not tested

Ø (a zero) as a hypothesis term indicates that the error term is to be used in a factor analysis

W as a hypothesis term indicates the constant term or the grand mean

W in a letter-number sequence indicates that the effect on the left is nested in the effect or cells on the right

a digit is required after FACTOR to indicate the type of factor analysis, see Table II

after FACTOR and its digit indicates that rotation of the factors is to be done: a digit must follow this to indicate type of rotation, see Table III

. (a period) must end the statement

ANOVA denotes an analysis of variance model, multivariate or univariate; CORREL denotes a correlation model, product moment, multiple or canonical; SERCOR denotes a serial correlation model, biserial or canonical; FACTOR denotes a factor analysis model, including principal components; INTABL denotes interaction tables for analysis of variance; VARED denotes a variance reduction study of correlation and analysis of variance problems.

c. Variable Format Cards (required)

These cards describe the data records and are the usual variable format cards of FORTRAN. Ten cards are allowed. (There is no fixed format data option in this program.) It is required that cell identification be read in before the scores on the continuous variables.

7. Data (required)

The data may be on cards or a computer kept file (see columns 17-18 of the PROB card). If the data are on cards, put en   ↄ x cards after the data to make a blank, but complete, data record. If the data are on a tape or disk file, the end-of-file mark will signify the end of the data.

8. Individual Significance Test Control Cards (optional)

These cards control the use of expository calculations on designated tests of the model. At most 6 of the significance tests of the model can be subject to these techniques.

Cols	Information
1-2	The number of the significance test to which this information applies (obtained by counting equal signs from the right [or the period])
3-4	Probability statement times 100 for battery reduction, dimension reduction and rotation of canonical variates in ANOVA and CORREL

Cols	Information
6	Type of battery reduction procedure for error variables 1 for Efroymson's stepwise addition/elimination procedure 2 for Wherry-Doolittle stepwise addition procedure 3 for Step-up variable elimination procedure
7-8	Number of factors to be obtained in a factor analysis. For components analysis, the number of components is always the number of variables and for principal factor analysis this is always one less than the number of variables; no entries are required for these solutions. All other solutions require an entry here.
10	Factor analysis tests Blank for none 1 for Rao's test in canonical factor analysis 2 for Rippe's test
12	Type of rotation technique (see Table III)
14	1 for direct rotation of canonical variates; all factor analysis 2 for indirect rotation of canonical variates
16	1 for taxonomy of variables 2 for taxonomy of groups
17-18	Power of 0.1 (multiplied by 10) for the convergence criterion in uniqueness iterations of canonical factor analysis (see Table IV). Entering 16 produces a criterion of $0.1^{1.6} = .0251189$
20	1 if squared communalities are supplied for factor analysis
22	1 if weights are supplied for rotations of factors or canonical variates

9. Individual Significance Test Card Options (optional)

a. Communalities for factor analysis. FORMAT (8F10.0)

Squared communalities are entered in 10 column fields, 8 per card, for as many cards as necessary.

b. Weights for rotation of factors and canonical variates. FORMAT (8F10.0)

Weights are entered in 10 column fields. For each canonical variate or factor there must be as many weights as there are variables. The weights for each canonical variate must begin on a new card. When rotating canonical variates in ANOVA or CORREL, there must be as many sets of weights as there are canonical variates; if the rotation is a taxonomy of groups, each set of weights will have only as many weights as there are groups. For principal components, weights must be supplied for all components (as many as there are variables). For all forms of factor analysis, weights must be supplied for as many factors as are indicated in columns 7 and 8 of the individual significance test card or, if this is zero, as many as the number of variables.

10. TITLE Cards (optional)

TITLE cards may be used here to identify reanalyses.

11. Reanalysis Card (optional)

Cols	
1-4	ANLY (not ANALY)
6	1 if a new significance test card is used
8	number of new contrast card sets (must be used for changing the partitioning of the continuous variables)

The following features are the same as on the PROB card and appear in the same card columns.

10	Number of individual significance test cards for this analysis
12	Print cell means and variances (1 means yes)
14	Print reduced model matrix (1 means yes) Data file number not necessary

Cols	
20	Print only estimates or raw regression coefficients (1 means yes)
22	Controls printed output (Ø gives minimal output). (See Table I)
24	Controls printed output (Ø adds nothing to the above). (See Table I)
26	Type of covariance adjustment (Ø or 1 does the classical adjustment, 2 uses "error" weights for reduction)

All other control cards and their options remain the same as for the original analysis.

12. Several Problems

Many problems may be submitted with a single run.

13. FINISH card

Cols	
1-6	FINISH

Table 1

Printed Output Controls

Information which is printed when column 22 and column 24 of the PROB or ANLY card are coded.

Codes: Col 22	Codes: Col 24	Information	Cross-index for Table IX
		Correlation	
b,1,2*	---**	Correlation among the hypothesis variables	a
b,2	---	Standard deviation of hypothesis variables	-
b,1,2	---	Cross-correlations between hypothesis and error variables	b
2	3	Raw score weights for regressing hypothesis variables onto error variables	c
2	3	Standard score weights for regressing hypothesis variables onto error variables	d
2	1,3	Standard error about the regression line(s)	
b,1,2	---	Correlations among the error variables	e
b,2	---	Standard deviations of the error variables	
1,2	---	Regression sums of squares when total sums of squares of error variables are unities[1]	f
---	---	Statistical Summary	
---	---	Dimension Reduction Statistics (Canonical Correlation only)	

*Punching a blank (b), a 1 (one) or a 2 (two) forces printing of the information.
**This designates that col 24 does not control printing of this information.
Numbered footnotes can be found at the end of the table.

Codes		Information	
Col 22	Col 24		
b,2	---	Univariate F statistics for error variables[2]	g
b,2	---	Correlations between the variables and the canonical variates within the error set	h
1,2	---	Standard score weights for regressing variables onto canonical variates within the error set	i
2	1	Raw score weights for regressing variables onto canonical variates within the error set	j
2	2	Weights for regressing raw score variables onto error canonical variates with unit sums of squares	k
1,2	---	Regression sums of squares when total sums of squares of hypothesis variables are unities	1
b,2	---	Univariate F statistics for hypothesis variables[2]	
b,2	---	Correlations between the variables and the canonical variates within the hypothesis set	
b,2	---	Standard score weights for regressing variables onto canonical variates within the hypothesis set	
2	1	Raw score weights for regressing variables onto canonical variates within the hypothesis set	
2	2	Weights for regressing raw score hypothesis variables onto hypothesis canonical variates with unit sums of squares	
3	1,2,3	Print only the information indicated in column 24	

Analysis of Variance

b,2	---	Raw estimates of effects	m
1,2	---	Standardized estimates of effects[3]	o
2	1	Orthogonalized estimates of effects[4]	n
b,2	---	Univariate standard errors	

...ies		Information	
Col 22	Col 24		
1,2	---	Error dispersions reduced to correlations[5]	p
1,2	---	Hypothesis sums of squares when error sums of squares are unities[6]	f
2	3	Hypothesis sums of squares and error sums of squares (univariate)	
---	---	Statistical summary	
---	---	Dimension reduction statistics (discriminant analysis and MANOVA only)	
b,2	---	Univariate F statistics	g
b,2	---	Correlations between the variables and the canonical variates	h
b,2	---	Correlations between the variables and the canonical variates weighted by the square roots of the associated canonical variance[7]	q
2	2	Discriminant function coefficients for standard scores	r
2	3	Discriminant function coefficients for raw scores	j
b,1,2	---	Estimates of effects for the canonical variates[8]	s
2	4	Transformation matrix for obtaining canonical contrasts (unreduced)[9]	t
		Factor Analysis	
---	---	Correlation matrix or covariance matrix (as appropriate)	
---	---	Estimates of standard deviations[10]	
---	---	Communality estimates, when applicable	
---	---	Final communality estimates (canonical factor analysis only)	
1	---	Characteristic roots and vectors	

Codes		
Col 22	Col 24	Information
---	---	Factor coefficients (unrotated)
---	---	Images and anti images (image analysis only)
---	---	Canonical correlations (Canonical factor analysis only)
1	---	Regression weights for factor scores (total variance procedures only)
---	---	Statistical tests (if used)
1	---	Residuals from the correlation matrix
1	---	Mean and standard deviation of residuals (below the diagonal)
1	---	Frequency distribution of residuals (if program is set up to handle 20 or more continuous variables)
		Variance Reduction
1	---	Sum of squares among design parameters and percentage loss
1	---	Correlation among design parameters before reduction
1	---	Correlation among design parameters after reduction
1	---	Sum of squares among contrasts and percentage loss
1	---	Correlations among contrasts before reduction
1	---	Correlations among contrasts after reduction
1	---	Correlations among hypothesis sums of squares before reduction
1	---	Correlations among hypothesis sums of squares after reduction
---	---	Data on trace of hypothesis sum-of-squares matrix
---	---	Hypothesis variance data for individual variables and percentage loss

Codes		Information	Cross index for Table IX
Col 22	Col 24		
		Serial Correlation	
2	1	Category variables cross products matrix	-
0,2	---	Category frequencies for nominal contrasts	-
1,2	---	Correlations among category variables	a
1,2	1	Standard deviation of category variables	-
1,2	---	Cross correlation between category and error variables	b
0,1,2	---	Estimates of contrasts among category means (pooled across samples)	m
1,2	---	Standardized estimates of contrasts among category means (pooled across samples)	n
0,2	---	Correlations among the error variables	e
0,2	---	Standard deviations of the error variables	
1,2	---	Regression sums-of-squares when total sums-of-squares of error variables are unities	f
---		Statistical summary	-
---		Dimension reduction statistics	-
0,2	---	Univariate F statistics for error variables	g
0,2	---	Correlation between the variates and the canonical variates within the error set	h
2	3	Standard score weights for regressing variables onto canonical variates within the error set	i
2	3	Raw score weights for regressing variables onto canonical variates within the error set	j

Codes		Information	
Col 22	Col 24		
2	3	Weights for regressing raw score error variables onto error canonical variates with unit sums of squares	k
0,1,2	---	Estimates of effects for canonical contrasts among categories	s
1,2	---	Regression sums-of-squares when total sums-of-squares of category variables are unities	-
0,2	---	Univariate F statistics for category variables	-
0,2	---	Correlations between the variates and the canonical variates within the category set	-
2	3	Standard score weights for regressing within the category set	-
2	3	Raw score weights for regressing variates onto canonical variables within the category set (transformation to canonical contrasts)	-
2	3	Weights for regressing category variables onto canonical variates with unit sums-of-squares	-

Footnotes for Table I

1. Regression sums of squares when the total sums of squares of the error (hypothesis) variables are unities. This matrix has on its diagonal the squared multiple correlations between the error (hypothesis) variables and all hypothesis (error) variables. The off-diagonal elements become the correlations among the regressed variables when they are divided by the square roots of the diagonal elements.

2. Univariate F statistics for error (hypothesis) variables. These are the F tests of the multiple correlations for each error (hypothesis) variable regressed on all the hypothesis (error) variables.

3. Standardized estimates are the raw estimates divided by their standard errors. A standardized estimate of +1.0 is one standard error above the grand mean of the data. These figures are easy to relate to confidence intervals abo the grand mean.

4. Orthogonalized estimates are appropriate in nonorthogonal designs. They are what is left of the raw estimates after the nonorthogonality of the design has been accounted for. The analysis which is produced is an analysis of these estimates. Comparison of the raw estimates with the orthogonalized estimates is sometimes useful in determining the effects of nonorthogonality on the analysis. In E. Cramer's MANOVA these are the "Estimates."

5. Error dispersions reduced to correlations. In a multivariate analysis of variance the error term is a variance-covariance matrix. This is that error term reduced to a correlation matrix.

6. Hypothesis sums of squares when error sums of squares are unities. The diagonals of this matrix when multiplied and divided by the degrees of freedom give the univariate F-ratios. The off-diagonal entries are a type of "covariance F" and reflect the relationships among treatment effects on the variables.

7. Correlations between the variables and the canonical variates weighted by the square roots of the associated canonical variance. These values are related to "Student's" t: when the analysis is of two samples and a single variable, it is "Student's" t. In a MANOVA or discriminant analysis, summing the squares of these values for one variable across the canonical variates will produce the univariate F ratio for that variable.

8. Estimate of effects for the canonical variates. These are the mean discriminant scores when the grand mean is zero. These estimates always add to zero for each canonical variate (or discriminant variable).

9. Transformation matrix for obtaining canonical contrasts. This matrix, when used to multiply the contrasts, produces the canonical contrasts. In orthogonal designs it produces exactly the canonical contrasts; in nonorthogonal designs it produces the canonical contrasts ignoring adjustments for the lack of orthogonality.

10. Estimates of standard deviations. This program always assumes that the data are a sample and not a population. These are not standard deviations but sample estimates. The estimates are residual to any covariates or sampling design.

Table II

Factor Analysis Codes

Codes	Type of Factor Analysis
1	Principal components of dispersion
2	Principal components of correlation
3	Principal factor analysis
4	Image analysis
5	Canonical factor analysis
6	Maximum likelihood factor analysis
7	Alpha factor analysis
8	Minres factor analysis

Table III

Factor Canonical Variate Rotation Codes

Codes	Type of Rotation
1	Quartimax
2	Varimax
3	Equamax
4	Promax
5	Multiple groups
6	Orthogonal centroids
7	Orthogonal bounds
8	Orthogonal Procrustes

Table IV

Power of .01 for Convergence Criterion

Power x 10	Criterion
11	.0794329
12	.0630957
13	.0501188
14	.0398107
15	.0316227
16	.0251189
17	.0199526
18	.0158490
19	.0125893
20	.0100000

Note: It so happens that the criterion is modular in the second digit of the power of 0.1 while the first digit determines the number of zeros after the decimal; i.e., use of 24 gives a criterion of .00398107 and 44 gives a criterion of .0000398107.

Control Cards in Detail

1. TITLE Card

Cols	
1-4	TITL these letters exactly
5-80	Any description that can be keypunched
	As many TITLE cards may be used as the user sees fit. All cards must have the letters TITL in columns 1 to 4

2. PROBLEM Card

Cols	Codes	Explanation
1-4		PROB these letters exactly
5-6	NFMC	Number of variable format cards used (10 or fewer)
		The variable format cards are used to describe the scores as they appear on the data file. This information tells how many variable format cards are necessary to describe the format of the scores.
8	NCONT	Number of contrast card sets (at least 2; at most 7)
		Each contrast card set describes a factor of an analysis of variance design except that one additional set is used to describe the continuous variables of the linear model. Thus a one-way analysis of variance model must have two contrast card sets: one for the design and one for the continuous variables. A simple correlation problem must have a dummy contrast card set even if it is only one sample, and one set for the continuous variables.
		In many problems the contrast card sets will have only one card per set.

Cols	Codes	Explanation
10	INFO28	Number of individual significance test control cards (at most 6) For any tests calculated it is possible to enter an individual significance test control card to control use of any of several expository techniques such as rotation, battery reduction or the like. At most 6 tests may use these techniques, so that at most 6 individual significance test control cards may be used.
12	MPRINT	Punching a 1 (one) forces printing of the cell means and variance; leaving a blank avoids calculation and printing
14	KPRINT	Punching a 1 (one) forces printing of the reduced model matrix
17-18	INFILE	If the data are on cards and submitted with the set-up cards, leave blank. If the data are on a computer kept file, punch in the file number
20	INFO21	Punch a 1 (one) when only estimates of effects or regression coefficients are desired for <u>all</u> the tests. Otherwise leave blank
22	INFO9	This column controls what output is to be printed. Leaving it blank prints the bare essentials of the test. Punching a 1 (one) or 2 (two) prints additional output. See Table I
24	INFO1Ø	This column also controls what output is to be printed. See Table I
26	INFO16	This column controls regression in analysis of covariance and <u>must</u> be used for all covariate adjustments of data. Punching a 1 (one) produces the classical analysis of covariance adjustment. Punching a 2 (two) produces a reduction of the hypothesis sum of squares by the error regression weights. Failure to enter a code results in the classical adjustments.

Cols	Codes	Explanation
28	INE035	Punching a 1 in this column prevents copying of the data file for reanalysis. This is particularly useful for large data sets when no reanalysis is to be done; it saves time on the machine.

3. Contrast Card Sets

The description below shows that the required information may spread over more than one card, hence the term "card set" rather than "card."

One set is required for the continuous variables and one set for each factor of an ANOVA design. If the data comprise a single sample as often happens in factor analysis or correlation analysis, it is still necessary to include a contrast card set for a "factor of a design"; that is, treat the data as if it were a one-way design with a single level. For this it may be convenient to have a 1 at some position in each data record; or any constant digit can be used and recoded, even a blank. The order of the card sets must be the same as the order in which the data factor codes appear on the variable format statement.

It is not necessary that all "factors" of a design appear on the significance test card. This makes it possible to use a "factor" index code for file editing or in later reanalyses of the data. Use of a factor code for editing should be limited since the culled cases are rejected and listed on the output.

Cols	Name	Explanation
1	NTABLE	For a factor of an ANOVA design: Any letter except V or W, but no two factors may use the same letter For the continuous variables: The letter V only
2-3	ICONT	Contrast codes for design variables = b Regular contrasts: Also called nominal and deviation contrasts Example: for three levels of a factor mean 1/3 1/3 1/3 df_1 2/3 -1/3 -1/3 df_2 -1/3 2/3 -1/3 = 1 Special contrasts. Any set of contrasts the user wants. The matrix must be fed in as described below and must follow this card set immediately

Cols	Name	Explanation

= 2 Orthogonal polynomials (integral values only). This option generates orthogonal polynomial contrasts and design parameters as integral values. The number of levels is limited to 19 for numerical reasons. The metric of the polynomial (i.e., the linear coefficients) must be supplied and must follow this card set immediately.

= 3 Reverse Helmert contrasts: also called difference contrasts. These contrasts are as follows:

mean	1/4	1/4	1/4	1/4
df_1	-1	1	0	0
df_2	-1/2	-1/2	1	0
df_3	-1/3	-1/3	-1/3	1

It should be noted that the fractions involved in these contrasts are not exactly representable on computers. This may cause serious numerical problems due to rounding errors when the number of data cases is large, say a couple of hundred

= 4 Special design parameters. Any set of design parameters the user wants. The matrix must be fed in as described below and must follow this card set immediately.

= 5 Orthogonal polynomials (decimal values only). This option generates orthogonal polynomials with normalized coefficients. Coefficients can be generated up to order 39, and the polynomials of order 39 are accurate at least up to degree 5. The accuracy of the higher order polynomials is not guaranteed. The metric (i.e., the linear coefficients) must be supplied and must follow this card set immediately

For power functions of continuous variables

= Ø no powers generated

= N (after V). When the contrast card set describes the continuous variables, this number describes the number of powers of each variable that will be generated to form a new variable; that is, if n = 3, the variables V, V^2 and V^3 will be available for analysis. It will also be assumed that there are N times the number of variables (i.e., 3 x V) in the new analysis and that they are in the order V, V^2, V^3.

Cols	Name	Explanation
4-6	NLEV	Number of levels for the factor or, after V, the number of variables (not including powers)
8	KRDNA	= b if the dummy or, after V, continuous variables are to be labeled by number = 1 if the dummy or, after V, continuous variables are to be labeled by names which will be supplied immediately after this card set
10	KLINT	For making linear transformations of the continuous variables of an analysis. Three options are available. = 1 ordinary linear transformation. Coefficients are expected in the original order of variables. = 2 Potthoff and Roy transformation with the weighting matrix. The weighting matrix precedes the matrix of the transformations. = 3 Potthoff and Roy transformation with the inverse of the weighting matrix. The inverse of the weighting matrix precedes the matrix of the transformations.
11,12 and 2 column fields	LEVSUB	When the dummy or the continuous variables are to be subdivided into partitions, this is number of variables in each partition. The number of degrees of freedom or variables are punched in 2 column fields and must account for all degrees of freedom or the total number of variables. If no partitioning is done, ignore
the next 2 column field		,, (two commas). This is necessary whether or not the variables are partitioned. If the variables are not partitioned, the commas go into columns 11 and 12
in 2 column fields	RECODE	Recoding of factor level identification or reordering of the variables For recoding the level identification, enter the code as it appears in the data file in the order in which the codes are to be renumbered. For example, if the codes are 15, 5 and 31 and these are to be recoded as 3, 1 and 2 make the entries in columns 11 through 16 as b53115. Use two column fields throughout and account for all levels. Zero is an admissible level code.

Cols	Name	Explanation
		For reordering continuous variables, enter the serial numbers of the variables, as they appear on the data file, in the new order. Use two column fields throughout and account for the total number of variables. When several operations are done to the variables in sequence, the sequence of operations is 1. raising to powers 2. making linear transformations 3. reordering 4. partitioning 5. naming In employing these features, the user must keep in mind such problems as (1) linear transformations must include all powers of the variables, (2) partitioning takes place on transformed and reordered variables etc. When variables are raised to powers, columns 4 to 6 must contain the number of original variables, but all subsequent operations must take account of the original variables and their powers. That is, if there are 5 original variables and these are raised to the ~~third~~ power, the program will expect 15 linear transformations, 15 variables to reorder, 15 variables to partition and 15 names
in the next 2 column field		,, (two commas). This is necessary whether or not the variables are partitioned and whether or not reordering or recoding has been used. If neither feature is used, the commas go into columns 13 and 14
the next 20 or less columns	ANAME	Any alphameric description of the factor or variable set. This may be as long as 20 characters

Cols	Name	Explanation
Last Column		. (a period)

4. Contrast Card Set Options

These descriptions govern use of options found on the contrast cards. Of the 4 options, use of one precludes use of the others except for the variable names option. When variable names are supplied, the names cards must precede cards used for other options. All options requiring cards must follow the contrast card set to which they apply and precede any successive contrast card set.

a. Variable names cards

The use of this option enables printing of names on the output. Both continuous variables and design parameters (or contrasts) may be named, although dummy parameters for ANOVA interactions will always be numbered.

Names will have 8 characters (including blanks) and 10 names may be put on a card. When dummy parameters are named there must be as many names as there are degrees of freedom. When continuous variables are named there must be as many names as there are variables. When the continuous variables are reordered, the names must be supplied in the new order as the names are not reordered. When linear transformations are made, the names will be affixed to the transformed variables (including analysis of covariance). When powers of the variables are generated, the program expects as many names as there are variables times powers.

b. Special contrast or special design cards

This option enables use of special one-way contrasts and design parameters. They are entered as square matrices with as many rows and columns as there are levels of the factor. The first contrast or design parameter entered is always that for the constant term or grand mean and is the first row of the matrix. The weights for obtaining the constant term or grand mean need not be all equal nor do the contrast coefficients or design parameters need to sum to zero, although this is usually desirable.

The elements of the matrix are punched in 10 column fields, 8 elements per card. Each row must begin on a new card and may continue on as many cards as necessary.

c. Orthogonal polynomial metric

The coefficients of the linear polynomial are called the metric and are usually (depending on the experiment) the integers from 1 to N indicating equal spacing of the levels of treatment. Equal spacing is not necessary, however, and the metric need not be successive integers.

The coefficients of the linear polynomial are entered in 10 column fields, 8 per card and continued on as many cards as necessary.

d. Linear transformation of the continuous variables and Potthoff and Roy models.

This option allows for testing linear combinations of the continuous variables. To use option 1 it is necessary to use all the variables in the regression equations and to have as many regression equations as there are variables even though some of the coefficients might be zero or one. The regression equations are entered as a transformation matrix, with the equations as rows. Each row begins on a new card, the coeffici ts are punched in 10 column fields, 8 per card, and on as many c ds as necessary.

When using option 2 or 3 for Potthoff and Roy models, the weighting matrix or its inverse precedes the matrix of the transformations. The elements weighting matrix or its inverse are entered a row at a time, punched in 10 column fields, 8 per card the same as the transformation matrix. It is essential that the weighting matrix precede the transformation matrix.

When reordering is used in conjunction with linear transformations, the reordering takes place on the transformed scores.

5. Significance Test Card

The significance test card describes the model under which the data are to be analyzed and tested. Linear models in three styles of calculation are available: analysis of variance, correlation, and factor analysis. The calculations to be made are indicated by the acronyms ANOVA, CORREL, or FACTOR in the first few columns of the significance test card. Calculating interaction tables in a factorial sampling design can be indicated by the acronym INTABL. Variance reduction studies can be calculated by using the acronym VARED.

Use of the acronym ANOVA invokes solution of the linear model as an analysis of variance problem. Either one variable or many may be analyzed in the model. The null hypothesis test utilizes Wilks' lambda criterion as approximated by the F distribution. Although the lambda distribution is ostensibly multivariate, its degenerate cases, univariate ANOVA, Hotelling's T^2, Mahalanobis' distance D, discriminant analysis and "Student's" t are all handled automatically.

Use of the acronym CORREL invokes solution of the model as a correlation problem. The null hypothesis test is Wilks' lambda criterion as is the case for ANOVA problems. Again, the degenerate cases of canonical correlation: multiple correlation, product moment correlation, biserial correlation and point biserial correlation are all handled automatically.

Use of the acronym SERCOR invokes a serial correlation analysis model. Serial correlation is used to analyze characteristics of observations <u>within</u> samples (in contrast to ANOVA which analyzes differences <u>between</u> samples). Serial correlation models are a mixture of ANOVA and CORREL models. To calculate a biserial correlation between a dichotomous characteristic, A, and variables, V, the model statement

SERCOR:W=0;A=V.

is appropriate. This states that the effect for the grand mean, W, is swept out of A and V because W is to the left of the semicolon. The effect A is not swept out of the variables, V, so that A = V is analyzed as a correlation model. This is a simple model which could also be calculated as either an ANOVA or CORREL model.

More complex serial correlation models are common for a model with several samples, A, and one characteristic, B, common over the samples the following model is appropriate

SERCOR:W=0,A=0;B=V,AB=V .

Here W and A are to the left of the semicolon to remove the effects of grand mean and between sample variation from the data before analyzing the characteristic B. The effects for B are not removed from AB and V because they are to the right of the semicolon; likewise AB is not removed from V. The effect AB=V is a test to check the homogeneity of regression of B onto V in the various samples of A. Further discussion of the uses of serial correlation is found in the reference cited in the bibliography. Serial correlations can be performed with all the variations of the linear model: covariates, rotation of canonical variates, Potthoff and Roy models, dimension reduction, etc.

Use of the acronym <u>FACTOR</u> invokes a factor-analytic decomposition of the data. Several types of solutions are available as displayed in Table II. The acronym FACTOR <u>must</u> be immediately followed by a digit from Table II to denote which factor decomposition is to be used. It is possible to denote a rotation procedure by following this digit with a number sign (#) and another digit to denote type of rotation. When communality procedures other than squared multiple correlations are used, the individual significance test card also must be used. When less than all the factors are to be extracted, the individual significance test card must be used.

Use of the acronym INTABL generates interaction tables for ANOVA problems. Interaction tables are generated as a nesting procedure without including the grand mean in the model. Thus the statement ,WAB=V1, generates all the means of the variables in V for the AB interaction of an ANOVA model. It is possible to generate many interaction tables in the same model statement. The estimates used for constructing interaction tables are the raw estimates and are not subject to analysis of covariance adjustments. Interaction tables adjusted for covariates may be obtained by using regression transformations on the original variables.

Use of the acronym VARED generates a study of the sum of squares for hypothesis before and after reduction of the model to orthogonality. Model statements follow the rules for ANOVA and CORREL models. VARED studies are not subject to analysis of covariance adjustments except by using regression transformations on the original variables.

On the remainder of the card, special symbols, letters and numbers are used to designate the model to be analyzed.

Letters. The letters used are those from the contrast card set. That is, use of the letter A assumes that there is an effect and a factor of the design to be called A and that there is a contrast card set that has "A" punched in column 1.

There is also a contrast card set for the continuous variables which has a "V" in column 1. When the phrase ,A=V, is punched into the card, a sum of squares for the A effect will be generated and tested for its regression on the variables V.

If the design is a factorial, there may also be a B effect to test. For this we punch the phrase ,B=V, in the significance test card. To test the interaction effect we punch the phrase ,AB=V, in the card. The use of two or more letters adjacent generates the Kronecker product of the main effect design parameters necessary for testing interactions.

Order of Kronecker product. This program generates Kronecker products in the order determined by the order of the contrast card sets and ultimately by the input data record. Consider the term AB. If the contrast cards set for factor A precedes that for factor B (which is to say the identification code for factor A is to the left of the identification code for factor B on the data file) the AB dummy parameters are generated and the estimates printed in the order a_1b_1, a_2b_1, a_3b_1,..., a_nb_1, a_1b_2, a_2b_2,...,a_nb_2,a_1b_3, ..,a_nb_m. The same order is generated whether the interaction term is written BA or AB.

Equal sign. An equal sign (=) is used to separate the hypothesis variables from the error variables. The hypothesis variables are designated by letters and numbers to the left of the equal sign; the error variables are designated by letters and numbers to the right of the equal sign.

Commas, colons, and periods. Commas (,) are used to separate the tests from each other. A colon (:) is used to separate the model acronym from the tests; and a period (.) is always the last character in a model statement.

Grand mean and zero. The constant term or grand mean may be included in the model by using the phrase ,W=Ø, where W indicates the constant term as a hypothesis and Ø (a zero) indicates that there are no error variables for the hypothesis and no test to be made. The use of zero as an error term is the way in which a set of variables may be included in the model as if they were hypothesis variables but not tested. When Ø is used as a hypothesis, it indicates that the error term is to be subjected to factor analysis.

With this information it is possible to write a significance test card for a simple factorial analysis of variance:

(a) ANOVA:W=∅,A=V,B=V,AB=V.

or a significance test card for a factor analysis:

(b) FACTOR1:W=∅,∅=V.

Numbered partitions. The contrast card option for partitions makes it possible to subdivide the sets of design parameters and continuous variables into several subsets. The partitions are written on the significance test card as A1, A2, or V1, V3, B4, etc., the number referring naturally to the ordinal position of the partition (this use of numbers can be easily confused with the use of numbers in nested models).

With this information it is possible to write the following models and many others.

(c) Partitioned analysis of variance (such as orthogonal polynomial tests)

ANOVA:W=∅,A1=V,A =V,B=V,A1B=V,A2B=V.

(d) Correlation models

CORREL:W=∅,[illegible]=V.

(e) Homogeneity of regression in a one-way design

ANOVA:W=∅,A=∅,V1=∅,[illegible]=V[illegible].

(f) A principal factor analysis on several samples

FACTOR:W=∅,A=∅,∅=V.

(g) Correlation analysis from several samples

CORREL:W=∅,A=∅,V1=V.

Plus sign. The plus sign is used to pool several effects to produce a sum of squares. In an analysis of variance with several factors, one may pool higher order interactions and test them simultaneously for the nested variables V1+V2.

For instance,

(h) ,AB+AC+BC+ABC=V1+V .

When orthogonal polynomials are partitioned, this also allows re-pooling of high order polynomials for simultaneous testing, as follows,

(i) ,P1=V,P2=V,P3+P4+P5=V.

When + is used between two sets of variables, the variables are pooled before any other operation takes place: For example V1/V2+V3 indicates that both V2 and V3 are jointly covariates for V1.

When the parameters for a given effect are used both singly in some tests and pooled in other tests, the order in which the parameters are presented on the significance test card must always be the same for every test in which the parameters are used.

W and numbers for nested analyses. The letter W is commonly used in statistical literature to indicate nesting; BWA indicates that several samples, B, are nested in each of the levels or samples of A. This notation is expanded slightly here as follows.

(j) ,BWA1=V, indicates a test of B within the first level of A.

(k) ,BWA1+BWA2=V, indicates a pooled test of B within only the first two levels of A.

(l) ,BWA=V, indicates a test of B pooled for all levels of A.

(m) ,BWAC, indicates a test of B pooled over all cells of a two-factor design, AC.

(n) ,V1WA1=V2, indicates a test of the correlation between V1 and V2 within the first level of factor A.

Here, a number, used after a letter which follows a W, indicates the level, not a partition. A number used after a letter but before a W (or in the absence of a W) indicates a partition.

Slash (/) for analysis of covariance. The use of a slash indicates that all variates after the slash and before the next comma (or colon or period or minus) are to be covariates for the test indicated. A slash and variables immediately following the acronym indicates that all the tests in the model have those variables as covariates. This makes possible the following types of models.

(o) Analysis of covariance in a factorial analysis of variance

ANOVA:W=Ø,A=V1/V2,B=V1/V2,AB=V1/V2. or

ANOVA/V2:W=Ø,A=V1,B=V1,AB=V1.

(p) Partial correlation

CORREL:W=Ø,V1=V2/V3.

(q) Principal components analysis with covariates

FACTOR1/V2:W=Ø,Ø=V1.

(r) ANOVA:W=Ø,A=V1,B=V1/V2,AB=V1/V2+V3.

Ampersand (&) for extension. The use of an ampersand indicates that all the variables after the ampersand and before the next comma (or colon or period or minus) are to be used as extension variables to the canonical variates of the test indicated. An ampersand immediately following the acronym indicates that extension is to be done to all the canonical variates in all the tests in the model statement. This makes it possible to write the following models.

(s) Discriminant analysis of V1 and A extended to V2

ANOVA:W=Ø,A=V1&V2. or

ANOVA&V2:W=Ø,A=V1.

(t) Canonical correlation between P and V1 extended to V2 and V3.

CORREL:W=Ø,B=V1&V2+V3.

(u) Factor analysis of V1 extended to V2.

FACTOR4:W=Ø,Ø=V1&V2.

Number sign (#) for rotation. The use of # after the acronym FACTOR and its digit indicates that factors are to be rotated. A digit must follow to denote which rotation scheme is to be used. Table III lists the types of rotations available. This makes it possible to write the model.

(v) Alpha factor analysis with equamax rotation

FACTOR7#3:W=Ø,Ø=V.

(Rotation of canonical variates in ANOVA or CORREL can be done only by using the individual significance test card.)

Asterisk (*) for components of variance. In many components of variance models, the error term for a test can be generated as a regression sum of squares. More simply, the sum of squares for hypothesis of one test may be the sum of squares for errors of another test. Therefore, to make it easy to calculate the sum of squares for error we use the notation A*V to denote the regression of continuous variables V on dummy parameters A to form an error term. This notation is allowed only on the right of an equal sign where errors are designated.

It should be noted that this program will not solve multivariate components of variance models where the number of continuous variables V is more than the number of dummy parameters A or degrees of freedom for error.

Minus (-) for dimension reduction. In multivariate models, it may be desirable to remove the significant canonical variates of one effect from the variables to be analyzed for another effect: to wit, the significant canonical variates of an ANOVA interaction removed from consideration in a test of a main effect. If A and B are main effects in an ANOVA design and AB their interaction the statement

,A=V-AB=V.

will generate a test for A independent of any significant interaction effects.

Semicolon (;) for limiting sweeps in serial correlation. In ANOVA and CORREL models all hypothesis terms in the model statement are swept out sequentially from left to right leaving a residual sum-of-squares of the variables. In serial correlation, the hypothesis terms which lie to the left of the semicolon are swept out but those hypothesis terms to the right of the semicolon are not swept. The assumption being that hypothesis statements to the right of the semicolon refer to characteristics of the observations in the samples or populations.

Special notes.

(a) When effects are partitioned and repooled (for example V1+V3) the pooling must always be presented on the significance test card in the same order. That is, you cannot state A=V1+V3 and B=V3+V1 in the same model. You must state A=V1+V3 and B=V1+V3. Also, when effects are partitioned, they must stay partitioned: i.e., A=V is not an alternative for A1+A2=V when A has been partitioned.

(b) In one model statement, a set of parameters may be used only for a hypothesis, an error, an extension or a covariate but not for two of these.

6. Variable Format Cards

These cards describe the way in which the data on the observations appear on the data file. There may be as many as 10 cards to describe the format. The use of the variable format follows the customary FORTRAN restrictions.

Suppose the data appeared as follows:

Cols. 1-5	information to be ignored
6	level number of the first factor
7	information to be ignored
8-9	level identification of the second factor
10-13	a datum on the first continuous variable
14-18	a datum on the second continuous variable
19-21	information to be ignored
22-25	a datum on the third continuous variable
26-to end	information to be ignored

The information on these records can be read with the following variable format statement

(5X,I1,1X,I2,F4.0,F5.0,3X,F4.0)

It will be noticed that all identification numbers are expected in fixed (I, integer) format while all continuous variable scores are expected in floating (F, decimal) format.

About factor identification. The reading of factor level identification must precede the reading of variable scores. In the above example the data record is arranged so that this occurs naturally. It is possible that the factor level identification is interspersed among the variable scores; in this case the "tab" feature of format statements may be used to read the factor level identification before the variable scores. If this cannot be done the data will need to be rearranged to put level codes first in the records.

All records must contain factor level information. When single samples are analyzed, it will be necessary either to include a constant on the file or to fake a factor with one level. This may be done by reading a number off the record and recoding it to 1 or by reading a blank on the record and recoding it to 1.

On IBM machines it will also be necessary to note the following comment about reading integer variable scores in F format: for example reading the score 12 as F2.0.

The score 12 read as F2.0 occupies 2 characters. When the program copies this score to save it for reanalysis, it copies 12.0 in two characters 2.0 which does not include the 1: The copy is overflowed. When the copy file is read for reanalysis the program will register an overflow in subroutine DATVEC. To prevent this disability either manufacture the original file as 012 and read it as F3.0 or manufacture the original file as 12. and read it as F3.0.

7. Data

The data may be on any file as long as the file is designated on the problem card by a two-digit number in columns 17 and 18. Blanks in these columns indicate the data are on cards and follow the variable format statement.

If the data are on cards, a blank data record must follow the data: there must be as many blank cards as there are cards in the data record of one observation. If the data are on a tape or disk file an ordinary end-of-file mark will do.

As a rule it is simpler to arrange factor level identification codes first on the data record before the continuous variable scores because these codes must be read in first in the record. This expedient is not a necessity on many machines because of the "tab" feature of FORTRAN IV compilers.

8. Individual Significance Test Control Card

The individual significance test cards are used to control procedures which can be applied to particular significance tests and not to others. These procedures are expository in nature and not usually subject to statistical testing. Arbitrarily, the number of statistical tests in the model to which these procedures can be applied is limited to six.

Each individual significance test control card is presented with the options which apply to it. The control cards are presented in the reverse order from the order of the tests on the significance test card (i.e., last to first).

Column	Name	Information
1-2	NOW	The number of the significance test to which this information applies. To obtain this number from the significance test card start at the period and, proceeding backwards toward the acronym, count the number of equal signs until the test which applies is reached, use that number in columns 1 and 2.
3-4	CRIT	The probability, multiplied by 100, used to control (1) additions and deletions of variables in battery reduction, (2) dimension reduction (use of minus signs) in ANOVA and CORREL and (3) the number of canonical variates to rotate in ANOVA or CORREL.
6	--	Type of battery reduction procedure for error variables =1 Efroymson's stepwise addition/elimination procedure =2 Wherry-Doolittle stepwise addition procedure =3 Step-up variable deletion procedure
7-8	MAXFAC	Number of factors to be utilized in factor analysis tests, rotations and extensions. This does not apply to the two principal components solutions as they always obtain as many components as there are variables. It does not apply to principal factor solutions as it always obtains one less factor than there are variables. All other procedures require an entry.
10	--	Statistical test procedure for factor analysis =Ø none =1 Rao's test, for canonical factor analysis only =2 Rippe's test
12	--	Type of rotation technique for factor analysis or multivariate ANOVA or CORREL. (Use of this is not necessary if the # (number sign) is used after the acronym FACTOR.) See Table III for the usable codes.

Column	Name	Information
14	--	For direct or indirect rotations =1 Direct rotation of canonical variates or factor analysis rotation =2 Indirect rotation of canonical variates
16	--	For taxonomy of rotation =1 for taxonomy of variables =2 for taxonomy of groups (in MANOVA only)
17-18	--	Criterion for uniqueness covergence in canonical factor analysis. This criterion is entered as an exponent (multiplied by 10) of 0.1 so as to give a wide range of values to convergence. The entry may be any two-digit number (see Table IV). For example, the entry 31 produces a convergence criterion of $0.1^{3.1}$ = .000794329.
20	--	Communalities supplied for factor analysis =1 if squared communalities are supplied for factor analysis. Only three procedures use this option: principal factor analysis, canonical factor analysis and image analysis. If communalities are supplied, they must immediately follow this card. =Ø if squared communalities are to be the squared multiple correlations between the variable and all other variables. IT IS NOT NECESSARY to use an individual significance test control card to have squared multiple correlations used as communalities. All common variance factor analysis procedures have an automatic default to squared multiple correlation.
22	--	=1 if weights are supplied for multiple groups, orthogonal centroids or orthogonal bounds rotations. If weights are supplied, they must immediately follow this card or the supplied squared communalities.

9. Individual Significance Test Control Card Options

a. Supplied communalities

Communalities (squared) must be entered 8 per card, in 10 column fields, and in the order in which the variables are to be analyzed.

b. Weights for rotations

The weights on each variable which determine a transformation vector are entered 8 to a card on as many cards as necessary. Each weight is entered in a 10 column field. In each transformation vector there must be as many weights as there are variables and each transformation vector must begin on a new card.

The number of transformation vectors to be entered is determined by the problem. For ANOVA and CORREL problems, there must be as many vectors as there are canonical variates; i.e., the minimum of the number of variables (degrees of freedom) for the hypothesis or the number of error variables. For factor analysis the number of sets of weights must be either the number of variables or the number of factors specified in columns 7 and 8 of the individual significance test control card.

10. TITLE Cards (optional)

TITLE cards may be used to identify reanalyses.

11. Reanalysis Card (optional)

This card allows, in the same run, reanalysis of the data. Two major changes are allowed in reanalysis: (a) change of any or all contrast card sets and (b) change of the model statement. There are some features of the first analysis which cannot be changed on reanalysis: (a) the number of contrast card sets used cannot be increased (although some factors of the design may be ignored in the model statement); (b) the number of levels or variables may not be increased (although some levels or variables may be ignored on the model statement); and (c) the format of the data (the variable format card) may not be changed.

Cols	Information
1-4	ANLY, these letters exactly
6	If a new significance test card is used to alter the model analyzed, punch a 1 (one) in column 6. (Each change of model must be accompanied by a reanalysis card)
8	The number of changed contrast card sets which are to be used in reanalysis. This feature can be used to alter several features of the model. (a) The contrast used for a given "factor" of the design may be changed. The letter designation of a "factor" may not be changed. (b) The way in which the dummy parameters and continuous variables are partitioned into subsets may be changed. (c) The names of variables may be changed. (d) The transformations of variables may be included or changed. To change these features simply follow the original instructions above.

The following features are the same as described for the PROB card and appear in the same card columns.

10	The number of individual significance test control cards for this reanalysis
12	Print cell means and variances (1 means yes)
14	Print reduced model matrix (1 means yes)
20	Print only estimates or raw regression coefficients (1 means yes)
22	Controls printed output. See Table I
24	Controls printed output. See Table I
26	Type of covariance adjustment (0 or 1 does the classical adjustment, 2 uses the error regression weights for reduction)

The options of columns 10 to 26 must be reinstated for each reanalysis because doing a reanalysis erases the controls of the previous analysis.

12. Several Problems

Many sets of data may be analyzed in one run. The cards for each set may follow each other.

13. FINISH Card

A card with FINISH in columns 1-6 may follow the last problem.

Programming Notes

The VARAN program is written in FORTRAN IV for an IBM 360-65 computer. Insofar as possible it is written to be compatible with any computer which uses FORTRAN-like compilers. Such features as may need changing are accessible without much labor. The following list shows some of these features.

1. All multiple precision statements are titled "DOUBLE PRECISION."
2. All input-output units are designated by integer constants which can be changed in the main program.
3. All dimensioning constants for changing the size of arrays are centralized in the main program and four executive subroutines. This allows for easy changing of size parameters and adaption of the program to special problems. The arrays and their dimensioning constants are listed in Tables VI through VIII. The program is distributed in three sizes, Tiny, Standard, and Large. The FORTRAN decks are the standard size while instructions are on the tape for implementing the tiny and large sizes.

The program is put together in two basic sections, one for the clerical part of setting up the model, and another for doing the mathematics of the solution. The clerical section is controlled by the executive subroutines SUBEX1, SUBEX3 and SUBEX4. The mathematical section is controlled by the executive subroutine SUBEX2 with the factor analysis subsection controlled by SUBEX5. Table V lists each of the subroutines with a brief description of its purpose. This table lists each major executive subroutine and after it the miscellaneous subroutines used by it as well as the major subroutines which are called by the executive subroutine. The order of the listing is roughly dictated by the following considerations:

1. the order in which the program executes
2. the order in which the overlay structure is put together, and
3. the order in which the FORTRAN decks are listed on the mailer tape (the exception here is that the executive subroutines are at the head of the mailer tape)

Comment cards are used quite liberally to explain the functions of subroutines. At the head of each subroutine there are several comment cards which describe the function of the subroutine and how it operates.

Accuracy of calculations. The programming sequence was arranged in such a way as to be convenient and accurate for multivariate problems. Therefore some of the univariate side statistics have been manipulated excessively outside the main flow of calculations and do not serve as checks for numerical accuracy. Accuracy checks should be made on the multivariate calculations.

Assembly language subroutines. Two of the subroutines MOVCHR and BINBCD are written in assembler code because FORTRAN IV does not perform the necessary operations. Another IFRMA is word size dependent. Listings of these subroutines are provided from the mailer tape but brief descriptions are also included here.

MOVCHR. The calling statement is: Call MOVCHR(NS,SA,NR,RA). The routine moves the NSth character in SA to the NRth position in SA. In the calling sequence NS is the integer position of the character to be moved, SA is the address of the source vector; NR is the integer position to receive the character, RA is the address to the receiving vector.

EXAMPLE:

Dimension A(4), B(4)

where A contains ABCDEFGHIJKLMNOP and B contains QRSTUVWXYZ123456

Call MOVCHR(12,A,3,B)

Result: A unchanged

B contains QRLTUVWXYZ123456

BINBCD. This is a real function which takes an integer, I, and converts it to alpha format, right justified. The following example is given in IBM hexadecimal notation.

Example X = BINBCD(I,10)

Input	Output
I = 00000001	X = 404040F1
I = 00000006	X = 404040F6

Although the authors are not certain (having fallible memories) to the best of our knowledge VARAN uses only a single precision space filling option; this is why the second argument is 10 in VARAN. The routine as originally devised is generalized to allow general integer conversion, zero fill, space fill, insertion of implied decimal point, and uses single or double precision results.

IFRMA. This is a function which converts integral numbers in ALPHA format to integral numbers in INTEGER format. Zero is returned if a space is input

Examples in HEX(360) code:

K = IFRMA(X)

Input	Output
X = F1404040	K = 00000001
X = F6404040	K = 00000006
X = 40404040	K = 00000000

Table V

Subroutine Name	Function	Called from
MAIN	Main program	
MATA	Add matrix A to matrix B and put in C	MANY
MATX	Multiply matrix A times matrix B and put in C	MANY
ERROR	Writes out error messages	MANY
IMPINV	Subroutine to obtain an improved inverse	MANY
INVT	Inversion of asymmetric matrix-Gauss pivotal	IMPINV
IFRMA	Function to convert alphabetic characters to integers	MANY
MOVCHR	Moves a character	MANY
LOOKUP	Search table for value equal to X	MANY
BINBCD	Converts binary to EBCDIC (IBM 360)	RDSIGT
SUBEX1	Executive subroutine to set up data vectors	MAIN
RDCONT	Read contrast cards and contrast card options and set up one way design parameter matrices	SUBEX1
DEVDES	Obtains deviation contrast design parameters	RDCONT
SPCON	Reads in special contrasts (transposed) and computes design variables for special contrasts	RDCONT
OPTPOL	Computes orthogonal polynomial design variables	RDCONT
HELDES	Obtains reverse Helmert design parameters	RDCONT

Table V (Cont'd)

Subroutine Name / Function	Called from
SPDES Reads in special one-way design matrices	RDCONT
POTROY Reads in weighting matrices and transformations for Potthoff and Roy models	RDCONT
RDSIGT Read the significance test card for the first time and set up the parameter codes for the terms listed. The scanning is done from left to right stopping at the period	SUBEX1
DATVEC Set up the data vector for each data case	SUBEX1
NEST Generates an identity matrix for nested effects	DATVEC
REMOD Print reduced model matrix. Generate data vectors for cells which are present	SUBEX1
WSTAT Finds the cell means and STD DEVS	SUBEX1
SUBEX3 Executive subroutine to obtain cross products matrix and sweep it	MAIN
SCP Accumulate and store cross-product matrices of the data vectors	SUBEX3
SWEEP Perform Gaussian sweep by matrix blocks	SUBEX3
SUBEX4 Scan the significance test card and set up for each individual test. First scan up to the colon to get permanent covariates, extensions and rotation types which hold for all the tests. Then skip to the period and scan backwards one test at a time, between commas and minus signs.	MAIN
HYERCO Set up matrices for a hypothesis-error only, hypothesis-cov, hypothesis-extension, or hypothesis-vector product analysis	SUBEX4
NOHYCO Set up matrices for no hypothesis-error, no hypothesis-covariate, no hypothesis-extension, or no hypothesis vector product analysis	SUBEX4
INTER Set up matrices for interaction tables	SUBEX4

Table V (Cont'd)

Subroutine Name	Function	Called from
APLB	Form matrices for the analysis subroutines	MANY
RECON	Reconstruction of matrices to conform to each other in terms of what has been swept	APLB
GETM	Get the I,J matrix from the SWP tape	APLB
SUBEXQ	Subroutine for controlling analytic sequence	MAIN
ATOBT	Subroutine to move matrix A to B and transpose it	MANY
MAB	Multiply matrix A times matrix B and put in C	MANY
MABT	Multiply matrix A times matrix B-transposed and put in C	MANY
MATB	Multiply matrix A-transpose times matrix B and put in C	MANY
OUTAB	Subroutine to print out tables	MANY
PROB	Subroutine to get probabilities in several variance distributions including F, t, normal deviates and chi-square	MANY
UNDFLW	Subroutine to centralize IBM ERRSET controls	MANY
WILKS	Calculate the overall F-test for Wilks lambda criterion, for multiple roots calculate chi-square for Bartlett's dimensionality reduction	MANY
HOW	Subroutine to obtain eigenvectors and values of a matrix	MANY
TRIDI	Tri-diagonalization subroutine DWM 1517-UB	HOW

Table V (Cont'd)

Subroutine Name	Function	Called from
EIGVEC	Set up simultaneous equations for eigenvector with eigenvalue E	HOW
EIGVAL	Eigenvalue subroutine for tri-diagonal matrices DWM 1517-UB	HOW
MEANS	Subroutine to provide estimates only or interaction tables	SUBEX2
DIMRED	Subroutine to perform dimension reduction between separate hypotheses of ANOVA	SUBEX2
ADJHYP	Subroutine to adjust hypothesis variables sum of squares and hypothesis-errors cross products	SUBEX2
ADJERR	Subroutine to adjust error sum of squares and raw estimates for covariates	SUBEX2
REDHYP	Subroutine to reduce hypothesis-errors cross products by error regression weights	SUBEX2
ESTADJ	Subroutine to obtain raw estimates adjusted for covariates from raw cross products	SUBEX2
ANAL	Performs calculations for ANOVA and CORREL	SUBEX2
DECOMP	Subroutine to decompose a symmetric matrix A into TT' where T is triangular	ANAL
PRANOV	Subroutine to print ANOVA calculations from ANAL	SUBEX2
PRCOR	Subroutine to print correlation calculations from ANAL	SUBEX2
PRSER	Subroutine to print serial correlation calculations from ANAL	SUBEX2
SUBEX5	Subroutine to sort out factor analysis procedures	SUBEX2
TRANV	Subroutine to follow immediately after DETER	MANY

Table V (Cont'd)

Subroutine Name	Function	Called from
DETER	Obtain determinant of A	MANY
SIMEQ	Solve (A + PHI*I)X = B . A is symmetric .	MANY
RESOL	Solve. (A + PHI*I)V = X subject to V(TR)*V = CON (usually = 1.) Restricted least squares.	MANY
FIRFAC	Subroutine to generate first five solutions of factor analysis	SUBEX5
ALPHA	Subroutine for alpha factor analysis: Michael Browne, circa 1968	SUBEX5
MALFA1	Maximum likelihood factor analysis. Version 2. Variables giving Heywood cases are partialled out.	SUBEX5
MINRES	Least squares factor analysis. Fixed unit weights.	SUBEX5
PRIFAC	Subroutine to print factor analysis results	SUBEX5
RAO	Subroutine to compute Rao's test of significance for canonical factor analysis	PRIFAC
RIPPE	Subroutine to compute Rippe's test for completeness of factorization	PRIFAC
ROTCON	Subroutine for controlling rotations	SUBEX2
VARMAX	Subroutine to perform quartimax, varimax, and equamax rotations	SUBEX2
PROMAX	Performs promax rotations	SUBEX2
WATCEN	Subroutine for doing three rotations, multiple groups, orthogonal centroids and orthogonal bounds	SUBEX2

Table V (Cont'd)

Subroutine Name.	Function	Called from
GRAM	Factor to lower triangle and orthogonal matrix	WATCEN
PRIROT	Subroutine for printing results of rotations	SUBEX2
EXTEND	Subroutine to perform and test extensions	SUBEX2
BATRED	Subroutine to do battery reduction techniques on error variables	SUBEX2
ELIM	Subroutine to partial out variables from a canonical correlation problem	BATRED
VARED	Subroutine to calculate loss of variance due to orthogonalization of design or of covariance analysis	SUBEX2
INDSIG	Subroutine to read individual significance test cards and options	SUBEX4

Table VI

Arrays with Variable Subscripts

Array		Subroutines with Fixed Dimensions
VARNAM(MNL,MNF)	(double precision)	MAIN,SUBEX1,SUBEX4
KSTART(MNT)		MAIN,SUBEX1,SUBEX3,SUBEX4
TNAME(MNC,MNT)		MAIN,SUBEX4
QNAME(MNCQ)		MAIN,SUBEX4
X(MNSIGC*80)		MAIN,SUBEX4
HYPNAM(MNL)	(double precision)	MAIN,SUBEX2,SUBEX4
ERRNAM(MNL)	(double precision)	MAIN,SUBEX2,SUBEX4
OTHNAM(MNL)	(double precision)	MAIN,SUBEX2,SUBEX4
LEVSUM(MNL,MNF)=LEVSUB		MAIN,SUBEX1,SUBEX4
ILHE(MNHE)		MAIN,SUBEX4
DUM(MNL,MNL)		SUBEX1
CONT(MNL,MNL,MNF)		SUBEX1
NTABLE(MNF)		SUBEX1
ICONT(MNF)		SUBEX1
NLEV(MNF)		SUBEX1
CMPVAR(MNF)		SUBEX1
RDNAM(MNF)		SUBEX1
TRAN(MNL,MNL)		SUBEX1
NPART(MNF)		SUBEX1
LEVCUM(MNL,MNF)		SUBEX1
RECODE(MNL,MNF)		SUBEX1
ANAME(MNA,MNF)		SUBEX1
IS(MNF)		SUBEX1
IE(MNF)		SUBEX1
IC(MNF)		SUBEX1
IN(MNF)		SUBEX1
NAFCD(MNF,MNT)		SUBEX1
ICELL(MNF)		SUBEX1
VAR(MNL)		SUBEX1
IVEC(MNCD)		SUBEX1
IFL(MNF)		SUBEX1
KSTART(MNT)		SUBEX1
FMT(NFMT)		SUBEX1
VECT(MNL)		SUBEX1
[illegible](3,MISV,MICELL)		SUBEX1
IDENT(MNCELL,MNF-2)		SUBEX1
IRM(MNCELL)		SUBEX1
A(NORD,NORD)		SUBEX2
B(NORD,NORD)		SUBEX2
C(NORD,NORD)		SUBEX2
D(NORD,NORD)		SUBEX2
E(NORD,NORD)		SUBEX2

Table VI (Cont'd)

Arrays with Variable Subscripts

Array	Subroutines with Fixed Dimensions
V1(NORD)	SUBEX2
V2(NORD)	SUBEX2
V3(NORD)	SUBEX2
V4(NORD)	SUBEX2
V5(NORD)	SUBEX2
AHYP(MNCQ) also AHYP(NORD)	SUBEX2
DVEC(MNT)	SUBEX3
E(MAXD)	SUBEX3
IBEG(MNT)	SUBEX3
N1(MNT)	SUBEX3
N2(MNT)	SUBEX3
K1(MNT)	SUBEX3
K2(MNT)	SUBEX3
A(MNL,MNL)	SUBEX3
B(MNL,MNL)	SUBEX3
C(MNL,MNL)	SUBEX3
D(MNL,MNL)	SUBEX3
NROW(MNT)	SUBEX3
V(MNL)	SUBEX3
LHE(MNHE,5)	SUBEX4
A(MNL,MNL)	SUBEX4
B(MNL,MNL)	SUBEX4
C(MNL,2*MNL)	SUBEX4
X(MNSIGC*80)	SUBEX4
PNAME(MNC)	SUBEX4
YD(MNL)	SUBEX4
IROWC(MNHE)	SUBEX4
D(MNL,MNL)	SUBEX4
AINTG(MNL)	SUBEX4
BDY(MNL-30)	SUBEX4
TOTNAM(MNL) (double precision)	SUBEX4

Table VII

Dimensioning Constants

MNL=NORD= Maximum number of levels for a factor and one more than the maximum number of continuous variables
Originates in MAIN, SUBEX1, SUBEX2, SUBEX3, and SUBEX4
Standard version=40

MNF= Maximum number of factors (including continuous variables) plus 1
Originates in MAIN and SUBEX1
Standard version=8

MNCD=MNT=NM= Maximum number of entries in the data vector; maximum number of terms on the significance test card; or maximum number of matrices in core at one time (in SUBEX3)
Originates in MAIN, SUBEX1, and SUBEX3
Standard version=258=2**MNF+2

MNC= Maximum number of words in each term name on the significance test card
Originates in MAIN, SUBEX1 and SUBEX4
Standard version=3=12 characters

MNCELL= Maximum number of cells for calculation of means and sigmas within cells
Originates in SUBEX1
Standard version=516=2**(MNF+1)+4

MAXD= Maximum dimension of a vector in SCP
Originates in SUBEX3
Standard version=20,000, greater than 4*MNL**2
(It is important that this be as large as possible without [illegible]tating the size of storage. It limits the size of the sum-of-cross-products matrix. When the SCP matrix is not large enough to hold the entire set of parameters, two or more passes are made in a loop which is the slowest part of the program. Making two or more passes through this loop greatly increases the cost of running the program.)

MNCQ= Maximum number of words in QNAME (a significance test)
Originates in SUBEX4
Standard version=40=160 characters
(Through an error, the printout will only list NORD words of the test statement.)

MNSIGC= Maximum number of significance test cards
Originates in SUBEX4
Standard version=9

Table VII (Cont'd)

Dimensioning Constants

NFMT= Maximum number of words in the variable format statement
Originates in SUBEX1
Standard version=200=10 cards

MLHE= Maximum number of items in each of the LHE lists
Originates in MAIN and SUBEX4
Standard version=20

MNA= Maximum number of characters in a factor name (the alphabetic
description on the contrast card)
Originates in SUBEX1
Standard version=20

MNSV= Maximum number of variables to be printed across a page for
cell means and standard deviations
Originates in SUBEX1
Standard version=8

Table VIII

Alterable Constants

Unnamed common in MAIN and standard version designations

KIN	=	System input unit = 5
KOUT	=	System printed output unit = 6
KPNCH	=	System punched output unit = 7
ISAVE	=	Temporary storage unit = 19
ITEMP	=	Temporary storage unit = 18
ISTOR	=	Temporary storage unit = 17
IKEEP	=	Temporary storage unit = 16
IBAG	=	Temporary storage unit = 15
ISIGT	=	Temporary storage unit = 14
IOTFIL	=	Temporary storage unit = 13
ISCPT	=	Temporary storage unit = 12
ISWPT	=	Temporary storage unit = 11
INFILE	=	Data input file if not KIN
UNDER	=	Smallest number machine will handle = 1.0E-70
OVER	=	Largest number machine will handle = 1.0E70
CNEARO	=	Constant for controlling rounding error = .000001
INFO	=	Array for controlling analytic sequence INFO(35) to INFO(50) unused on June 1, 1972

Table IX

Constants in INFO

```
INFO(1)  = 1 If analysis of variance
         = 2 If correlation
         = 3 If factor analysis
         = 5 If interaction tables
         = 6 If variance reduction
INFO(2)  = 1 If error sum of squares is residual
         = 2 If error sum of squares is generated from estimates and
             design parameters
INFO(3)  = ANOVA:  Degrees of freedom for hypothesis
           CORREL:  Number of variables in hypothesis set
INFO(4)  = Order of error sum of squares matrix
INFO(5)  = ANOVA:  Number of variables
           CORREL:  Number of variables in error set
INFO(6)  = Degrees of freedom for error
INFO(7) Controls battery reduction
INFO(7)  = 1 Efroymson's stepwise procedure
         = 3 Step-up procedure
         = 2 Wherry-Doolittle procedure
INFO(8)  = Probability level for adding and dropping variables in
           battery reduction
INFO(9) and INFO(10) Control standard analysis print-out
INFO(9)  = 0 Print minimal sets
         = 1 Print intermediate sets
         = 2 Print everything
         = 3 Print INFO(10) output only
INFO(10) For ANOVA
         = 1 Print orthogonalized raw estimates
         = 2 Print disc. fn. coef. for standard scores
         = 3 Print disc. fn. coef. for raw scores
         = 4 Print transf. matrix for canonical contrasts
INFO(10) For CORREL
         = 1 Print canonical raw score regression weights
         = 2 Print canonical raw score regression weights to obtain unit
           sum of squares regressed scores
         = 3 St. score and raw score weights for regressing hypothesis
           variables onto error variables
INFO(10) For factor analysis
         = 0 Print minimal set
         = 1 Print everything
INFO(10) For VARED
         = 0 Print data for hypothesis variables only
         = 1 Print data about design parameters, contrasts and
           correlations among hypothesis sums of squares
INFO(11) through INFO(14) Control rotation procedures
```

Table IX (Cont'd)

Constants in INFO

INFO(11) = Number of factors in factor analysis
= Criterion for determining number of significant canonical variates in MANOVA or canonical correlation rotation problems
INFO(12) = 1 If Quartimax
= 2 If Varimax
= 3 If Equamax
= 4 If Promax
= 5 If multiple groups
= 6 If orthogonal centroids
= 7 If orthogonal bounds
INFO(13) = 1 If factor analysis or direct rotation of canonical variates
= 2 If indirect rotation of canonical variates
INFO(14) = 1 If taxonomy of variables
= 2 If taxonomy of groups
INFO(15) = Number of covariates
INFO(16) = 1 Classical covariance adjustments
= 2 Covariance reduction by error regression weights
INFO(17) = Number of extension variables
INFO(18) = 1 If a test with dimension reduction
= 2 If a test with dimension reduction and for same
= 3 If a test without dimension reduction but for same
INFO(19) = Number of dimensions which have been set up for reduction
INFO(20) = Number of observations minus number of parameters swept out
INFO(21) = 1 If raw estimates only for ANOVA
INFO(22) = 1 Principal components of dispersion
= 2 Principal components of correlation
= 3 Principal factor analysis
= 4 Image analysis
= 5 Canonical factor analysis
= 6 Maximum likelihood
= 7 Alpha
= 8 Minres
INFO(23) = 0 No communalities required
= 1 If communalities start with S.M.C.
= 2 If communalities supplied
INFO(24) = Number of factors to be removed and tested = INFO(11)
INFO(25) = Significance tests procedure
= 0 If none
= 1 If Rao's
= 2 If Rippe's
INFO(26) = Blank
INFO(27) = Power of .1 to be used as a convergence criterion (multiplied by 10)
INFO(28) = Number of individual significance test cards
INFO(29) = Number of calls to INLSIG (i.e., number of the current hypothesis test)
INFO(30) = 1 If rotations are multiple groups, orthogonal centroids or orthogonal bounds
INFO(31) Counts the number of individual significance test cards read in INLSIG
INFO(32) = Current INFO(3)
INFO(33) = INFO(3) for previous test
INFO(34) = 1 For calls to Wilks from DIMRED
INFO(35) = 1 To prevent copying of data file for reanalysis

Mathematical Notes

These notes are to describe the mathematical calculations involved in solving the linear model. As the steps are described, the subroutines which execute these steps will be indicated in capital letters within parentheses.

The program is written as a full linear model. That is, as the vector of measurements for each observation are read into the computer, dummy parameters (or design parameters or fixed variables) are added to the vector according to the requirements of the linear model (DATVEC). The new data vector including dummy parameters and measurement variables is stored on tape or disk.

When all observations have been read in and data vectors have been constructed, a sum-of-squares and cross-products matrix is then constructed (...). In order to accommodate analyses which have a large number of parameters, the sum-of-squares and cross-products matrix is not necessarily stored entirely in the machine. The program generates and stores as large a portion of this matrix as possible, then rewinds itself and generates another large portion of the matrix. Unfortunately when it is necessary to generate the sum-of-squares and cross-products matrices in portions, the expense of a run increases greatly. (Fortunately this is a very rare occurrence.) The sum-of-squares matrix is partitioned into submatrices which are completely determined by the effects listed on the model card. This matrix is stored on disk as partitioned blocks and elements within blocks rather than elements within a matrix.

The complete sum-of-squares matrix is then reduced to the sums of squares and cross products for the various effects as required by the model (SWEEP). This is done by sweeping out the effects listed in the model in the order in which they appear (from left to right) on the significance list card. A statement of the order appears on the printout. The sweeping process is done as a straightforward Gaussian reduction, but instead of reducing the matrix to a diagonal an element at a time, the process reduces the matrix a partitioned block at a time. That is, the Gaussian sweep is generalized from an element at a time to a submatrix at a time. Otherwise, the process used is that described by Bock (1963).

Subsequent to the sweeping process, the appropriate sums-of-squares and cross-products matrices for each linear analysis are assembled (SUBEX4) and passed on to the appropriate analytic section.

ANOVA and CORREL Solutions

Analysis of variance and correlation problems stem from the determinantal equation

$$|A-\lambda B|=0 \qquad (1)$$

where A is a matrix of the sum of squares and cross products for hypothesis and B is, for ANOVA, the sum of squares and cross products for errors or, for CORREL, the sum of squares and cross products for "totals" (using the convention that the sum of squares for errors plus the sum of squares for hypothesis is the sum of squares for totals regardless of what other parameters are in the model for the data).

The solution of this equation (ANAL) is accomplished as follows.

1. The matrix B is scaled to correlations and A scaled compatibly.

2. The scaled B is decomposed (DECOMP) into a product of triangular matrices. This triangular matrix is then inverted and pre- and postmultiplied onto the scaled A. This results in reducing the original determinantal equation into a single matrix eigenvalue problem.

3. The eigenvalues are calculated and eigenvectors determined (HOW).

This description is the essential flow of the calculations. However, there are many side statistics calculated to provide expository material for the analyses. Table X lists the side statistics and gives their formulations.

In Table X the symbols have the following meanings.

Z-the hypothesis variables: in ANOVA the dummy parameters, dummy variables design parameters, fixed variables or independent variables; in correlation the independent variables or predictors. They are assumed to have mean zero.

X-the error variables: in ANOVA the observed variables, the dependent variables, the continuous variables: in correlation the criteria or dependent variables. They are assumed to have mean zero.

When the determinantal equation (1) is written for correlation problems using Z and X, it can be written in matrix notation as

$$|(X'Z)\ (Z'X)^{-1}\ (Z'X) - \lambda(X'X)| = 0. \qquad (2)$$

The values for λ are the squared canonical correlations. When the equation is written for analysis of variance, the matrix B is

$$B = (X'X) - (X'Z)\ (Z'Z)^{-1}\ (Z'X)$$

which for convenience of writing will still be notated $X'X$ even though it is a residual error or within-cells matrix. The values of λ here become ratios of sums of squares.

Z'Z A of equation 1

[illegible] is the diagonal matrix of the square roots of the diagonal of Z'Z. In correlation problems this is divided by a constant to give the standard deviations of the Z variables.

σ_X is the diagonal matrix of the square roots of the diagonal of $X'X$. In correlations problems this is divided by a constant to give the standard deviations of the X. In analysis of variance this is divided by a constant to give the standard errors of the X. (The matrix $X'X$ is not the same in both solutions.)

Q is the triangular decomposition of $\sigma_X^{-1} X'X\sigma_X^{-1}$

T is the matrix of eigenvectors of

$$|Q'^{-1}\sigma_X^{-1}(X'Z)(Z'Z)^{-1}\ (Z'X)\ \sigma_X^{-1}Q^{-1} - \lambda I| = 0$$

DFH degrees of freedom for hypothesis

DFE degrees of freedom for error

The final matrix equation of the solution can be written

$$\lambda^{-1/2}TQ'^{-1}\sigma_X^{-1}(X'Z)(Z'Z)^{-1}(Z'X)\sigma_X^{-1}Q^{-1}T'\lambda^{-1/2} = I.$$

The various parts of this formulation are entered in Table X and cross-referenced to Table I to help the reader recognize the printed output if he is not familiar with the style of nomenclature used here.

Factor Analysis Solutions

The factor analytic solutions available in VARAN are common ones and the nomenclature used is standard. The subroutines for doing the maximum likelihood, ALPHA and MINRES procedures, were written by Michael Browne and were merely adapted for use here. These routines print out information which is not generally used but was retained nevertheless.

Interaction Tables

The mathematics of calculation of interaction tables is simple enough. The parameters for nesting within the cells of the desired interaction are generated, and the estimates for this design are calculated, viz: WAR

Table X

Formulation and Nomenclature

Formula	Correlation name[1]	ANOVA name[1]	[illegible] Table 1
$X'Z(Z'Z)^{-1}$	-	either raw or orthogonalized estimates	m,n
$\sigma_X^{-1}X'Z(Z'Z)^{-1}$	-	standardized estimates	c
$\sigma_X^{-1}X'Z\sigma_Z^{-1}$	cross correlations	-	b
$X'Z(Z'Z)^{-1}$	raw score regression weights	-	c
$\sigma_Z^{-1}Z'Z\sigma_Z^{-1}$	hypothesis variable correlations	design parameter correlations	a
$\sigma_X^{-1}X'Z(Z'Z)^{-1}$	standard score regression weights	-	d
$\sigma_X^{-1}X'Z(Z'Z)^{-1}Z'X\sigma_X^{-1}$	hypothesis sums of squares when total sums of squares are unities	hypothesis sums of squares when error sums of squares are unities	f
diag $(\sigma_X^{-1}X'Z(Z'Z)^{-1}Z'X\sigma_X^{-1})$DFH:DFE	univariate F ratios	univariate F ratios	g
$TQ'^{-1}\sigma_X^{-1}X'Z(Z'Z)^{-1}$	raw score weights for canonical variate scores	discriminant function coefficients	j
$TQ'^{-1}\sigma_X^{-1}X'Z(Z'Z)^{-1}\sigma_Z^{-1}$	standard score weights for canonical variate scores	standardized discriminant function coefficients	i,r
$\Lambda^{-1/2}TQ'^{-1}\sigma_X^{-1}X'Z(Z'Z)^{-1}$	(see Table I)	(see Table I)	k
$TQ'^{-1}\sigma_X^{-1}X'Z(Z'Z)^{-1}Z'X$	-	estimates of effects for canonical variates (mean discriminant scores)	s
$(TQ'^{-1}\sigma_X^{-1}X'Z(Z'Z)^{-1}Z'X)^{-1}$	-	transformation matrix for obtaining canonical contrasts	t
QT	correlations between error variables and error canonical variates	correlations between variables and canonical variates	h
$QT\Lambda^{1/2}$	-	correlations between variables and canonical variates weighted by the square roots of the canonical variances	q
$\sigma_X^{-1}X'X\sigma_X^{-1}$	correlations among the[1] error variables	error dispersions reduced to[1] correlations	e,r

[1] $X'X$ means different things in ANOVA and in CORREL models. σ_X is the square root of the estimate of population variance: in ANOVA this is the standard error, in CORREL this is the standard deviation and in SERCOR the standard deviation in most models.

produces the estimates for the cells of the AB face of a factorial design. The grand mean must not be included in ANIAPL models.

The acronym INTAPL concludes the ANOVA calculations when estimates have been obtained, then prints them out.

Variance Reduction Analysis

Variance reduction analysis is a simple procedure for evaluating the effect of imbalance in an ANOVA design due to unequal cell sizes. Basically, the hypothesis sums of squares for the given test is generated both before and after taking account of the inequality of cell sizes. (The grand mean is not removed from the data.) The "before" sums of squares is assumed to be 100 per cent of the data available. When the "after" sums of squares is larger than the "before" a negative loss is indicated.

Dimension Reduction Analysis

This feature is available through the significance test card and is indicated by a minus sign (-). In use it pertains to a feature of multivariate analysis of variance and canonical correlation which has no univariate analog.

To explain its use, suppose we are given two independent statistical hypotheses, both involving the same error variables, such as an interaction and a main effects test in a MANOVA design of p variables. Suppose the interaction test, AB=V, has $r < p$ significant dimensions (or roots). These r dimensions of the p dimensions of error have the same properties as r variables in r univariate analyses and suggest that the main effects of these r variables need not be examined further. However, the remaining $p-r$ dimensions of error are not interactive and can be reasonably tested over the main effect A.

The test of the main effect of A on the p-r dimensions of error is reasonable and can be effected by obtaining p-r linear combinations of the error variables which are independent of the significant interaction discriminant variables and analyzing these.

The method of dimension reduction removes the significant interaction discriminant variables from the sums of squares for the test A=V. This forces those same linear combinations of the error variables to have zero roots in the determinantal equation $|A-\lambda B| = 0$, thus removing them from the analysis. The main effects test A=V is then calculated, and the multivariate probability tests are compiled as if there were only p-r variables in the test. The univariate F ratios are not altered.

The notation on the significance test card to produce this test of A is

ANOVA:W=0,A=V-AB=V.

with an individual significance test card used to indicate the probability level in the test of AB=V which determines both significance and the number of canonical variates to be removed from A=V.

In addition to this application to analysis of variance, dimension reduction can also be applied to canonical correlation,

CORREL:W=0,V1=V2-V3=V2.

and as a type of analysis of covariance,

ANOVA:W=0,A=V1-V2=V1.

Rotation Procedures for Canonical Variates and Factors

Eight rotation procedures are available for multivariate solutions to linear models. Four of these are familiar techniques of factor analysis: Quartimax, Varimax, Equamax and Promax. These four techniques apply mathematical criteria to the solutions and rotate the correlations between the canonical variates and the variables to a solution that is presumably more comprehensible to the researcher.

The remaining four rotation procedures apply external criteria to the rotational procedure. In multiple groups, orthogonal centroids and orthogonal bounds, a set of regression weights are entered and used to determine the correlations between the rotated canonical variates and the variables. This makes all these procedures Procrustian in form; however, they are not explicitly Procrustian in that they use regression weights, not target coefficient matrices. The multiple groups procedure obtains an oblique solution. The orthogonal centroids scheme is similar but uses the first set of weights optimally, then uses the second set of weights optimally but orthogonal to the first set. Successive weights are fitted optimally but orthogonal to all the previously fitted weights.

Orthogonal bounds is Procrustian in form also. First a multiple group oblique solution is found then an orthogonal envelope is found which has a least squares fit about the multiple group solution. The correlations between the canonical variates and the variables are rotated to the envelope.

The orthogonal Procrustes scheme is due to Norman Cliff and is referenced in the bibliography. The calculation scheme is due to Roger Pennell. Here the weights supplied are the "target" matrix of coefficients and not regression weights.

SAMPLE INPUT

```
//CFM$$502 JOB (6600,55,10,9,999),'04161,IRN,HALLCHAS',CLASS=E
//STEP1 EXEC PGM=VARAN,REGION=(,168K),TIME=5
//STEPLIB DD DSN=HALLLIB,UNIT=2314,VOL=SER=RESLIB,DISP=SHR
//FT05F001 DD DDNAME=SYSIN
//FT06F001 DD SYSOUT=A,DCB=(RECFM=FBA,LRECL=133,BLKSIZE=3458)
//FT11F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//FT12F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//FT13F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//FT14F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//FT15F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//FT16F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//FT17F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//FT18F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//FT19F001 DD SPACE=(CYL,(1,1)),DISP=NEW,UNIT=DISK
//SYSIN DD *
TITLE PROBLEM 6, TWO WAY FACTORIAL
TITLE MAXIMUM ANOVA PRINTOUT
TITLE CELL MEANS PRINTED
TITLE PRINT REDUCED MODEL MATRIX
TITLE TEST PROBLEM FROM HALL AND CRAMER
PROB 1 3    1 1        2
A     2      ,,,,FACTOR A.
B     2      ,,,,FACTOR B.
V      5 1   ,,,,CONTINUOUS VARIABLES.
ERROR 1 ERROR 2 ERROR 3 ERROR 4 ERROR 5 ERROR 6
ANOVA:N=0,A=V,B=V,AB=V.
(2I1,10X,6F6.0)
115   41    569.   156.   104.   506.   4.     9.
115   42    475.   120.   105.   360.   4.    16.
115   43    541.    85.    82.   815.   4.    16.
115   44    779.   104.   104.   331.   4.    17.
115   45    587.    98.    53.   564.   3.    13.
115   46    841.   129.    71.   519.   5.     5.
115   47    907.    90.    49.   416.   3.    16.
115   48    698.    76.    53.   442.   4.     9.
115   49    849.   132.    89.   459.   5.     1.
115   50    505.   166.   207.   474.   4.    -9.
126   51    557.    91.    62.   513.   3.    26.
126   52    643.   114.    52.   416.   4.    16.
126   53    714.    81.    50.   491.   4.     0.
126   54    611.   125.    80.   630.   3.    25.
126   55    713.    84.    57.   471.   4.     4.
126   56    644.    97.    63.   453.   4.     5.
126   57    533.    83.    55.   540.   4.     8.
126   58    536.   125.    85.   364.   3.    24.
126   59    988.   109.    94.   554.   4.     7.
126   60    554.   120.    91.   503.   4.     4.
218   71    455.    72.    67.   623.   4.    12.
218   72    846.    96.    79.   539.   4.     2.
218   73    704.   109.    45.   355.   2.     7.
218   74    953.   142.    67.   432.   3.    22.
218   75    553.    97.    80.   495.   3.    16.
```

```
218   76    592.  82.   67.   362.  2.    17.
218   77    529.  85.   50.   657.  2.    14.
218   78    556.  99.   67.   589.  1.    18.
218   79    419.  103.  77.   452.  3.    15.
218   80    598.  78.   59.   397.  4.    0.
229   81    662.  75.   48.   385.  4.    14.
229   82    668.  88.   48.   361.  2.    8.
229   83    519.  110.  48.   391.  3.    11.
229   84    449.  90.   66.   484.  4.    9.
229   85    647.  80.   56.   482.  4.    2.
229   86    589.  64.   44.   337.  3.    15.
229   87    846.  73.   60.   598.  4.    6.
229   88    748.  96.   65.   601.  4.    10.
229   89    763.  135.  92.   480.  2.    12.
229   90    578.  102.  65.   683.  3.    -8.

TITLE SPECIAL CONTRASTS, HOMOGENEITY OF REGRESSION, CANONICAL CORRELATION,
TITLE ROTATION OF CANONICAL VARIATES, NAMED CONTRASTS AND MAXIMUM
TITLE CORRELATION PRINTOUT
TITLE TEST PROBLEM FROM HALL AND CRAMER
PROB 1 2 1          2
A 1   4 1   ,,11122122,,SPECIAL CONTRASTS.
1ST-4TH 2ND-3RD LEFTOVER
1.0        1.0        1.0        1.0
1.0                              -1.0
           1.0        -1.0
1.0        -1.0       -1.0       1.0
V    6 1   3 3,,,,VARIABLES.
ERROR 1 ERROR 2 ERROR 3 ERROR 4 ERROR 5 ERROR 6
CURREL:W=0,A=0,V2=V1,AV2=V1.
(I2,10X,6F6.0)
115   41    569.  156.  104.  506.  4.    9.
115   42    475.  120.  105.  366.  4.    16.
115   43    641.  83.   82.   815.  4.    16.
115   44    779.  104.  104.  331.  4.    17.
115   45    587.  98.   53.   564.  3.    13.
115   46    841.  129.  71.   519.  5.    5.
115   47    907.  90.   49.   416.  3.    16.
115   48    698.  76.   53.   492.  4.    9.
115   49    849.  132.  89.   459.  5.    1.
115   50    505.  166.  207.  474.  4.    -9.
126   51    557.  91.   62.   513.  3.    26.
126   52    649.  114.  52.   416.  4.    16.
126   53    714.  81.   50.   491.  4.    0.
126   54    611.  125.  80.   630.  3.    25.
126   55    713.  84.   57.   471.  4.    4.
126   56    644.  97.   63.   453.  4.    5.
126   57    593.  83.   55.   546.  4.    8.
126   58    536.  125.  85.   364.  3.    24.
126   59    988.  109.  94.   554.  4.    7.
126   60    584.  120.  91.   503.  4.    4.
218   71    935.  72.   67.   623.  4.    12.
218   72    846.  96.   79.   539.  4.    2.
```

218	73	704.	109.	45.	355.	2.	7.
218	74	453.	142.	67.	432.	3.	22.
218	75	553.	97.	80.	495.	3.	16.
218	76	592.	82.	67.	362.	2.	17.
218	77	529.	85.	50.	657.	2.	14.
218	78	556.	99.	67.	589.	1.	18.
218	79	419.	103.	77.	452.	3.	15.
218	80	598.	78.	59.	397.	4.	0.
229	81	662.	75.	48.	385.	4.	14.
229	82	668.	88.	48.	361.	2.	8.
229	83	519.	110.	48.	391.	3.	11.
229	84	449.	90.	66.	484.	4.	9.
229	85	647.	80.	56.	482.	4.	2.
229	86	589.	64.	44.	337.	3.	15.
229	87	846.	73.	60.	598.	4.	6.
229	88	748.	96.	65.	601.	4.	10.
229	89	763.	135.	92.	480.	2.	12.
229	90	578.	102.	65.	683.	3.	-8.

275 2 1 1
FINISH
/*
//

SAMPLE OUTPUT

```
VARAN:  LINEAR MODEL VARIANCE ANALYSIS

FIRST EDITION:  JUNE 1, 1972

EDUCATIONAL TESTING SERVICE

PRINCETON, N.J.  08540


 PROBLEM  0, TWO WAY FACTORIAL
 MAXIMUM ANOVA PRINTOUT
 CELL MEANS PRINTED
 PRINT REDUCED MODEL MATRIX
 TEST PROBLEM FROM HALL AND CRAMER
FACTOR  A
    FACTOR A
    FACTOR A HAS    2 LEVELS
    DEVIATION CONTRASTS
FACTOR  B
    FACTOR  B
    FACTOR B HAS    2 LEVELS
    DEVIATION CONTRASTS
FACTOR  V
    CONTINUOUS VARIABLES
    FACTOR V HAS    6 LEVELS
    VARIABLES NAMED:  ERROR 1    ERROR 2    ERROR 3    ERROR 4    ERROR 5    ERROR 6

THE ORDER OF EFFECTS IS :
                                          A
                                          B
                                          AB
                                          V

DATA FILE FORMAT    (2I1,10X,6F6.0)
```

REDUCED MODEL MATRIX
PARAMETER

CELL				
1 1	1.00	1.00	1.00	1.00
2	1.00	1.00	-1.00	-1.00
2 1	1.00	-1.00	1.00	-1.00
2 2	1.00	-1.00	-1.00	1.00

MEANS AND STANDARD DEVIATIONS

CELL	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1 1						
MEAN	685.0999	115.4000	91.7000	494.2000	4.0000	[illegible]
STD DEV	153.1223	30.4967	46.0943	133.0731	0.6667	[illegible]
N	10.	10.	10.	10.	10.	10.
1 2						
MEAN	658.8999	102.9000	68.9000	[illegible]	3.7000	[illegible]
STD DEV	130.0965	17.7606	16.8685	74.6673	[illegible]	[illegible]
N	10.	10.	10.	10.	10.	10.
2 1						
MEAN	668.5000	96.3000	[illegible]	490.0000	[illegible]	[illegible]
STD DEV	183.5635	19.8889	[illegible]	[illegible]	[illegible]	[illegible]
N	10.	10.	10.	10.	10.	10.
2 2						
MEAN	646.8999	91.3000	[illegible]	[illegible]	[illegible]	[illegible]
STD DEV	119.0882	20.7689	[illegible]	[illegible]	[illegible]	[illegible]
N	10.	10.	10.	10.	10.	10.

MODEL ANALYSED- ANOVA:R=0,A=V,B=V,AB=V.

ANALYSIS OF DISPERSION

HYPOTHESIS TESTED AB=V

RAW ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	1.150	1.875	4.050	-2.450	0.200	-1.750
NEG. SUM	-1.150	-1.875	-4.050	2.450	-0.200	1.750

STANDARDIZED ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	0.00774	0.08237	0.15454	-0.02224	0.25701	-0.21496
NEG. SUM	-0.00774	-0.08237	-0.15454	0.02224	-0.25701	0.21496

ORTHOGONALIZED ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	1.150	1.875	4.050	-2.450	0.200	-1.750
NEG. SUM	-1.150	-1.875	-4.050	2.450	-0.200	1.750

UNIVARIATE STANDARD ERRORS OF ESTIMATION

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	146.532	22.762	26.208	110.182	0.778	8.141

ERROR DISPERSIONS REDUCED TO CORRELATIONS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
ERROR 1	1.00000	-0.06310	-0.17101	0.07991	0.29513	-0.09719
ERROR 2	-0.06310	1.00000	0.63302	-0.12186	-0.12169	-0.05830
ERROR 3	-0.17101	0.63302	1.00000	0.02084	0.07587	-0.26122
ERROR 4	0.07991	-0.12185	0.02084	1.00000	0.07418	-0.11655
ERROR 5	0.29513	-0.12168	0.07587	0.07418	1.00000	-0.44021
ERROR 6	-0.09719	-0.05830	-0.26122	-0.11655	-0.44021	1.00000

HYPOTHESIS SUMS-OF-SQUARES WHEN ERROR SUMS-OF-SQUARES ARE UNITIES

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
ERROR 1	0.000067	0.000709	0.001329	-0.000191	0.002211	-0.001849
ERROR 2	0.000709	0.007540	0.014144	-0.002035	0.023523	-0.019675
ERROR 3	0.001329	0.014144	0.026535	-0.003818	0.044130	-0.036910
ERROR 4	-0.000191	-0.002035	-0.003818	0.000549	-0.006350	0.005311
ERROR 5	0.002211	0.023523	0.044130	-0.006350	0.073393	-0.061385
ERROR 6	-0.001849	-0.019675	-0.036910	0.005311	-0.061385	0.051541

UNIVARIATE SUMS-OF-SQUARES FOR HYPOTHESIS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	52.9	140.6	656.1	240.1	1.6	1[illegible]2.6

UNIVARIATE SUMS-OF-SQUARES FOR ERRORS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	794221.9	18651.6	24726.2	437043.8	21.[illegible]	2386.0

STATISTICAL SUMMARY

F-RATIO FOR WILKS LAMBDA CRITERION	0.5569
NUMERATOR DEGREES OF FREEDOM	6.
DENOMINATOR DEGREES OF FREEDOM	31.0
PROBABILITY	0.7609

UNIVARIATE F STATISTICS, 1 AND 36 DEGREES OF FREEDOM

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
F-RATIO	0.00240	0.27142	0.95524	0.01978	2.64215	1.84828
PROBABILITY	0.9612	0.6056	0.3349	0.8889	0.1128	0.1824

CORRELATIONS BETWEEN THE VARIABLES AND THE CANONICAL VARIATES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	0.025
ERROR 2	0.264
ERROR 3	0.496

ERROR 4	-0.071
ERROR 5	0.825
ERROR 6	-0.690

CORRELATIONS BETWEEN THE VARIABLES AND THE CANONICAL VARIATES
WEIGHTED BY THE SQUARE ROOTS OF THE ASSOCIATED CANONICAL VARIANCES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	0.049
ERROR 2	0.521
ERROR 3	0.977
ERROR 4	-0.141
ERROR 5	1.625
ERROR 6	-1.360

DISCRIMINANT FUNCTION COEFFICIENTS FOR STANDARD SCORES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	-0.16719
ERROR 2	0.16850
ERROR 3	0.22192
ERROR 4	-0.13543
ERROR 5	0.74444
ERROR 6	-0.32672

DISCRIMINANT FUNCTION COEFFICIENTS FOR RAW SCORES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	-0.00113
ERROR 2	0.00740
ERROR 3	0.00847
ERROR 4	-0.00123
ERROR 5	0.95[illegible]64
ERROR 6	-0.04013

ESTIMATES OF EFFECTS FOR THE CANONICAL VARIATES

VARIABLE	CANONICAL VARIATE 1
1	0.311
NEG. SUM	-0.311

TRANSFORMATION MATRIX FOR OBTAINING CANONICAL CONTRASTS (UNREDUCED)

VARIABLE	CANONICAL CONTRAST 1
1	3.21079

ANALYSIS OF DISPERSION

HYPOTHESIS TESTED B=V

RAW ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	11.950	4.375	7.350	2.500	-0.050	0.450
NEG. SUM	-11.950	-4.375	-7.350	-2.500	0.050	-0.450

STANDARDIZED ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	0.08045	0.19221	0.28045	0.02269	-0.06425	0.05527
NEG. SUM	-0.08045	-0.19221	-0.28045	-0.02269	0.06425	-0.05527

ORTHOGONALIZED ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	11.950	4.375	7.350	2.500	-0.050	0.450
NEG. SUM	-11.950	-4.375	-7.350	-2.500	0.050	-0.450

UNIVARIATE STANDARD ERRORS OF ESTIMATION

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	148.532	22.762	26.208	110.182	0.778	8.141

ERROR DISPERSIONS REDUCED T CORRELATIONS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
ERROR 1	1.00000	-0.06310	-0.17101	0.07991	0.29513	-0.09719
ERROR 2	-0.06310	1.00000	0.63302	-0.12186	-0.12169	-0.05830
ERROR 3	-0.17101	0.63302	1.00000	0.02084	0.07587	-0.26122
ERROR 4	0.07991	-0.12185	0.02084	1.00000	0.07418	-0.11655
ERROR 5	0.29513	-0.12168	0.07587	0.07418	1.00000	-0.44021
ERROR 6	-0.09719	-0.05830	-0.26122	-0.11655	-0.44021	1.00000

HYPOTHESIS SUMS-OF-SQUARES WHEN ERROR SUMS-OF-SQUARES ARE UNITIES

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
ERROR 1	0.007192	0.017182	0.025071	0.002028	-0.005744	0.004941
ERROR 2	0.017182	0.041049	0.059895	0.004846	-0.013722	0.011805
ERROR 3	0.025071	0.059895	0.087393	0.007070	-0.020022	0.017224
ERROR 4	0.002028	0.004846	0.007070	0.000572	-0.001620	0.001394
ERROR 5	-0.005744	-0.013722	-0.020022	-0.001620	0.004587	-0.003946
ERROR 6	0.004941	0.011805	0.017224	0.001394	-0.003946	0.003395

UNIVARIATE SUMS-OF-SQUARES FOR HYPOTHESIS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	5712.1	765.6	2160.9	250.0	0.1	8.1

UNIVARIATE SUMS-OF-SQUARES FOR ERRORS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	794221.9	18651.6	24726.2	437043.6	21.8	2385.0

STATISTICAL SUMMARY

F-RATIO FOR WILKS LAMBDA CRITERION	0.7258
NUMERATOR DEGREES OF FREEDOM	6.
DENOMINATOR DEGREES OF FREEDOM	31.0
PROBABILITY	0.6322

UNIVARIATE F STATISTICS, 1 AND 36 DEGREES OF FREEDOM

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
F-RATIO	0.25891	1.47775	3.14615	0.02059	0.16513	0.12221
PROBABILITY	0.6140	0.2320	0.0840	0.8867	0.6869	0.7287

CORRELATIONS BETWEEN THE VARIABLES AND THE CANONICAL VARIATES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	0.226
ERROR 2	0.541
ERROR 3	0.789
ERROR 4	0.064
ERROR 5	-0.181
ERROR 6	0.155

CORRELATIONS BETWEEN THE VARIABLES AND THE CANONICAL VARIATES
WEIGHTED BY THE SQUARE ROOTS OF THE ASSOCIATED CANONICAL VARIANCES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	0.509
ERROR 2	1.216
ERROR 3	1.774
ERROR 4	0.144
ERROR 5	-0.406
ERROR 6	0.350

DISCRIMINANT FUNCTION COEFFICIENTS FOR STANDARD SCORES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	0.51001
ERROR 2	-0.09 5
ERROR 3	1.05218
ERROR 4	0.05121
ERROR 5	-0.26728
ERROR 6	0.36244

DISCRIMINANT FUNCTION COEFFICIENTS FOR RAW SCORES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	0.00343
ERROR 2	-0.00433
ERROR 3	0.04015
ERROR 4	0.00046
ERROR 5	-0.34347
ERROR 6	0.044[illegible]2

ESTIMATES OF EFFECTS FOR THE CANONICAL VARIATES

VARIABLE	CANONICAL VARIATE 1
1	0.356
EG. SUM	-0.356

TRANSFORMATION MATRIX FOR OBTAINING CANONICAL CONTRASTS (UNREDUCED)

VARIABLE	CANONICAL CONTRAST 1
1	2.81242

ANALYSIS OF DISPERSION

HYPOTHESIS TESTED A=V

RAW ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	7.150	7.675	8.900	4.500	0.400	0.250
NEG. SUM	-7.150	-7.675	-8.900	-4.500	-0.400	-0.250

STANDARDIZED ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	0.04814	0.33719	0.33960	0.04084	0.51402	0.03071
NEG. SUM	-0.04814	-0.33719	-0.33960	-0.04084	-0.51402	-0.03071

ORTHOGONALIZED ESTIMATES OF EFFECTS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
1	7.150	7.675	8.900	4.500	0.400	0.250
NEG. SUM	-7.150	-7.675	-8.900	-4.500	-0.400	-0.250

UNIVARIATE STANDARD ERRORS OF ESTIMATION

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	148.532	22.762	26.208	110.182	0.778	8.141

ERROR DISPERSIONS REDUCED TO CORRELATIONS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
ERROR 1	1.00000	-0.06310	-0.17101	0.07991	0.29513	-0.09719
ERROR 2	-0.06310	1.00000	0.63302	-0.12186	-0.12169	-0.05830
ERROR 3	-0.17101	0.63302	1.00000	0.02084	0.07587	-0.26122
ERROR 4	0.07991	-0.12185	0.02084	1.00000	0.07418	-0.11655
ERROR 5	0.29513	-0.12168	0.07587	0.07418	1.00000	-0.44021
ERROR 6	-0.09719	-0.05830	-0.26122	-0.11655	-0.44021	1.00000

HYPOTHESIS SUMS-OF-SQUARES WHEN ERROR SUMS-OF-SQUARES ARE UNITIES

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
ERROR 1	0.002575	0.018035	0.018164	0.002184	0.027493	0.001642
ERROR 2	0.018035	0.126328	0.127230	0.015301	0.192579	0.011505
ERROR 3	0.018164	0.127230	0.128159	0.015411	0.193954	0.011587
ERROR 4	0.002184	0.015301	0.015411	0.001853	0.023326	0.001394
ERROR 5	0.027493	0.192579	0.193954	0.023326	0.293573	0.017539
ERROR 6	0.001642	0.011505	0.011587	0.001394	0.017539	0.001048

UNIVARIATE SUMS-OF-SQUARES FOR HYPOTHESIS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	2044.9	2356.2	3168.4	810.0	6.4	2.5

UNIVARIATE SUMS-OF-SQUARES FOR ERRORS

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
	794221.	18651.6	24726.2	437043.8	21.8	2386.0

STATISTICAL SUMMARY

F-RATIO FOR WILKS LAMBDA CRITERION	3.2863
NUMERATOR DEGREES OF FREEDOM	6.
DENOMINATOR DEGREES OF FREEDOM	31.0
PROBABILITY	0.0128

UNIVARIATE F STATISTICS, 1 AND 36 DEGREES OF FREEDOM

VARIABLE	ERROR 1	ERROR 2	ERROR 3	ERROR 4	ERROR 5	ERROR 6
F-RATIO	0.09269	4.54782	4.61301	0.06672	10.56862	0.03772
PROBABILITY	0.7625	0.0399	0.0385	0.7976	0.0025	0.8471

CORRELATIONS BETWEEN THE VARIABLES AND THE CANONICAL VARIATES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	0.064
ERROR 2	0.446
ERROR 3	0.449

ERROR 4	0.154
ERROR 5	0.679
ERROR 6	0.041

CORRELATIONS BETWEEN THE VARIABLES AND THE CANONICAL VARIATES
WEIGHTED BY THE SQUARE ROOTS OF THE ASSOCIATED CANONICAL VARIANCES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	0.304
ERROR 2	2.135
ERROR 3	2.146
ERROR 4	0.256
ERROR 5	3.251
ERROR 6	0.194

DISCRIMINANT FUNCTION COEFFICIENTS FOR STANDARD SCORES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	-0.12271
ERROR 2	0.48656
ERROR 3	0.18766
ERROR 4	0.11019
ERROR 5	0.99816
ERROR 6	0.55829

DISCRIMINANT FUNCTION COEFFICIENTS FOR RAW SCORES

VARIABLE	CANONICAL VARIATE 1
ERROR 1	-0.00083
ERROR 2	0.02138
ERROR 3	0.00716
ERROR 4	0.00100
ERROR 5	1.28269
ERROR 6	0.06858

ESTIMATES OF EFFECTS FOR THE CANONICAL VARIATES

VARIABLE	CANONICAL VARIATE 1
1	0.757
NEG. SUM	-0.757

TRANSFORMATION MATRIX FOR OBTAINING CANONICAL CONTRASTS (UNREDUCED)

VARIABLE	CANONICAL CONTRAST 1
1	1.32170

TEST WITH NO ERROR NOT DONE
N=0

```
VARAN:  LINEAR MODEL VARIANCE ANALYSIS

FIRST EDITION:  JUNE 1, 1972

EDUCATIONAL TESTING SERVICE

PRINCETON, N.J.  08540


 SPECIAL CONTRASTS, HOMOGENEITY OF REGRESSION, CANONICAL CORRELATION,
 ROTATION OF CANONICAL VARIATES, NAMED CONTRASTS AND MAXIMUM
 CORRELATION PRINTOUT
 TEST PROBLEM FROM HALL AND CRAMER
FACTOR A
    SPECIAL CONTRASTS
    FACTOR A HAS    4 LEVELS
    RECODING OF LEVEL IDENTIFICATION
         11   12   21   22
    VARIABLES NAMED:  1ST-4TH   2ND-3RD   LEFTOVER
    DESIGN PARAMETERS FROM SPECIAL CONTRASTS
           0.250000  0.250000  0.250000  0.250000
           0.500000  0.0       0.0      -0.500000
           0.0       0.500000 -0.500000  0.0
           0.250000 -0.250000 -0.250000  0.250000
FACTOR V
    VARIABLES
    FACTOR V HAS    6 LEVELS
         V1 HAS     3  DF
         V2 HAS     3  DF
    VARIABLES NAMED:  ERROR 1   ERROR 2   ERROR 3   ERROR 4   ERROR 5   ERROR 6

THE ORDER OF EFFECTS IS :  W
                           A
                           V2
                           AV2
                           V1

DATA FILE FORMAT    (I2,10X,6F6.0)


MODEL ANALYSED-   CORREL:W=0,A=0,V2=V1,AV2=V1.
```

ANALYSIS OF CORRELATION

HYPOTHESIS TESTED AV2=V1

CORRELATIONS AMONG THE HYPOTHESIS VARIABLES

VARIABLE	1	2	3	4	5	6	7	8	9
1	1.00000	0.03990	0.12068	0.21596	-0.00756	-0.18165	-0.37714	-0.05031	0.28729
2	0.03990	1.00000	0.28467	-0.04526	-0.05162	0.04563	0.06850	0.08512	-0.00128
3	0.12068	0.28467	1.00000	-0.12172	0.01420	0.08638	0.23116	0.08418	-0.09980
4	0.21596	-0.04526	-0.12172	1.00000	-0.14949	-0.09821	-0.23435	0.05465	-0.14889
5	-0.00756	-0.05162	0.01420	-0.14949	1.00000	0.59461	-0.05252	-0.66212	0.14571
6	-0.18165	0.04563	0.08638	-0.09821	0.59461	1.00000	-0.21409	-0.03179	-0.42401
7	-0.37714	0.06850	0.23116	-0.23435	-0.05252	-0.21409	1.00000	-0.02555	0.24939
8	-0.05031	0.08512	0.08418	0.05465	-0.66213	-0.03179	-0.02555	1.00000	-0.24566
9	0.28729	-0.00128	-0.09980	-0.14889	0.14571	-0.42401	0.24939	-0.24566	1.00000

STANDARD DEVIATIONS OF HYPOTHESIS VARIABLES

VARIABLE	1	2	3	4	5	6	7	8
	44.60148	32.93964	27.25245	0.26450	0.25153	0.1567	7.08263	5.08575

VARIABLE	9
	2.03088

CROSS CORRELATIONS BETWEEN HYPOTHESIS AND ERROR VARIABLES

VARIABLE	1	2	3	4	5	6	7	8	9
ERROR 1	-0.12[illegible]	0.04753	-0.09421	0.14751	-0.02[illegible]	-0.220[illegible]	0.05887	-0.15267	0.14873
ERROR 2	-0.17[illegible]	0.00556	0.05685	0.31676	-0.09904	0.12957	-0.33984	0.07972	-0.50279
ERROR 3	-0.15027	0.00210	-0.09445	0.12611	-0.17499	0.03435	-0.42906	0.09965	-0.43401

STANDARD WEIGHTS FOR REGRESSION HYPOTHESIS VARIABLES ONTO ERROR VARIABLES

ERROR	HYPOTHESIS 1	2	3	4	5	6	7	8
ERROR 1	-1.05148	0.45454	0.08824	115.76407	27.76607	-192.41502	-6.49805	-6.311[illegible]
ERROR 2	-0.18429	0.05602	0.16226	23.28732	6.05812	-23.97102	-3.63197	-0.25156
ERROR 3	-0.18551	0.05022	0.05291	1.12867	-26.21429	-5.93063	-5.25415	-1.37488

ERROR	HYPOTHESIS 9
ERROR 1	11.20874
ERROR 2	-3.91291
ERROR 3	-2.94881

STANDARD SCORE WEIGHTS FOR REGRESSING HYPOTHESIS VARIABLES ONTO ERROR VARIABLES

ERROR	HYPOTHESIS 1	2	3	4	5	6	7	8
ERROR 1	-0.31772	0.09856	0.01627	0.20716	0.04725	-0.25615	-0.11706	-0.13175
ERROR 2	-0.35446	0.08092	0.19044	0.26527	0.06541	-0.20316	-0.41648	-0.03540
ERROR 3	-0.31355	0.03839	0.05463	0.01151	-0.24981	-0.04422	-0.53003	-0.16071

ERROR	HYPOTHESIS 9
ERROR 1	0.15401
ERROR 2	-0.34225
ERROR 3	-0.22669

STANDARD ERROR ABOUT THE REGRESSION LINE FOR ERROR VARIABLES

VARIABLE	ERROR 1	ERROR 2	ERROR 3
	158.34528	20.36580	24.10719

CORRELATIONS AMONG ERROR VARIABLES

VARIABLE	ERROR 1	ERROR 2	ERROR 3
ERROR 1	1.00000	-0.01600	-0.19987
ERROR 2	-0.01600	1.00000	0.64724
ERROR 3	-0.19987	0.64724	1.00000

STANDARD DEVIATIONS OF THE ERROR VARIABLES

VARIABLE	ERROR 1	ERROR 2	ERROR 3
	147.80696	23.22012	26.39468

REGRESSION SUMS-OF-SQUARES FROM TOTAL SQ S-OF-SQUARES OF ERROR VARIABLES ARE UNITIES

VARIABLE	ERROR 1	ERROR 2	ERROR 3
ERROR 1	0.16555	0.04270	0.01291
ERROR 2	0.04270	0.44054	0.36422
ERROR 3	0.01291	0.36422	0.59532

STATISTICAL SUMMARY

F-RATIO FOR WILKS LAMBDA CRITERION	1.0075
NUMERATOR DEGREES OF FREEDOM	27.
DENOMINATOR DEGREES OF FREEDOM	64.9
PROBABILITY	0.4755

DIMENSION REDUCTION STATISTICS FOR LARGE N. HERE, N= 24.

CANONICAL VARIATE	1	2	3
CANONICAL CORRELATIONS	0.7063	0.5967	0.3626
INDIVIDUAL LAMBDAS	0.995	0.189	0.172
CHI-SQUARE FOR ROOTS K THROUGH LAST	21.2114	6.8762	3.2845
DEGREES OF FREEDOM	27.	16.	7.
PROBABILITY	0.7751	0.9759	0.8571

UNIVARIATE F STATISTICS FOR ERROR VARIABLES, 9 AND 24 DEGREES OF FREEDOM

VARIABLE	ERROR 1	ERROR 2	ERROR 3
MULTIPLE CORRELATION	0.4066	0.6637	0.6272
F-RATIO	0.52819	2.09981	1.72866
PROBABILITY	0.8396	0.0711	0.1368
ALIENATION	0.91361	0.74797	0.77889

CORRELATIONS BETWEEN THE VARIABLES AND THE CANONICAL VARIATES WITHIN THE ERROR SET

	CANONICAL VARIATE		
VARIABLE	1	2	3
ERROR 1	0.1499	0.9307	0.5183
ERROR 2	0.9121	-0.2678	0.3103
ERROR 3	0.8364	-0.1714	-0.5207

STANDARD SCORE WEIGHTS FOR REGRESSING VARIABLES ONTO CANONICAL VARIATES WITHIN THE ERROR SET

	CANONICAL VARIATE		
VARIABLE	1	2	3
ERROR 1	0.26092	0.99483	0.09164
ERROR 2	0.58721	-0.46405	1.09606
ERROR 3	0.50847	0.32779	-1.21175

RAW SCORE WEIGHTS FOR REGRESSION VARIABLES WITH CANONICAL VARIATES WITHIN THE ERROR SET

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 1	0.00177	0.04673	0.00062
ERROR 2	0.02529	-0.01998	0.04720
ERROR 3	0.01926	0.01242	-0.04591

WEIGHTS FOR REGRESSION RAW SCORE VARIABLES TO FORM CANONICAL VARIATES WITH UNIT SUMS-OF-SQUARES

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 1	0.00210	0.01036	0.06100
ERROR 2	0.03009	-0.03165	0.07631
ERROR 3	0.02292	0.01967	-0.07422

REGRESSION SUMS-OF-SQUARES WHEN TOTAL SUMS-OF-SQUARES OF HYPOTHESIS VARIABLES ARE UNITIES

VARIABLE	1	2	3	4	5	6	7	8	9
1	0.05392	-0.01411	0.01903	-0.06932	0.03354	0.01025	0.07491	0.00340	0.07875
2	-0.01411	0.00839	0.00765	0.03117	-0.00110	0.00017	-0.00830	-0.00325	-0.02137
3	0.01903	0.00765	0.05262	0.01789	0.02823	0.03898	0.03005	0.01337	-0.01934
4	-0.06932	0.03117	0.01789	0.12974	-0.02445	0.01176	-0.08146	0.00048	-0.13110
5	0.03354	-0.00110	0.02823	-0.02445	0.05490	0.00123	0.07452	-0.00825	0.06370
6	0.01025	0.00017	0.03898	0.01176	0.00123	0.06774	-0.04077	0.04200	-0.08977
7	0.07491	-0.00830	0.03005	-0.08145	0.07452	-0.04077	0.19057	-0.04225	0.20828
8	0.00340	-0.00325	0.01557	0.00048	-0.00826	0.04200	-0.04225	0.02557	-0.06429
9	0.07875	-0.02137	-0.01934	-0.13110	0.06370	-0.08977	0.20828	-0.06429	0.27490

UNIVARIATE F STATISTICS FOR HYPOTHESIS VARIABLES, [illegible] 30 DEGREES OF FREEDOM

VARIABLE	1	2	3	4	5	6	7	8
MULTIPLE CORRELATION	0.2322	0.0916	0.2294	0.3602	0.1[illegible]	0.2603	0.4370	0.[illegible]
F-RATIO	0.56994	0.08465	0.55541	1.49084	0.36166	0.72660	2.36044	[illegible]
PROBABILITY	0.6391	0.9679	0.6485	0.2370	0.7811	0.5441	0.0912	0.[illegible]
ALIENATION	0.97267	0.99579	0.97334	0.93288	-0.99239	0.96553	0.[illegible]	0.[illegible]

VARIABLE	9
MULTIPLE CORRELATION	0.5338
F-RATIO	3.98408
PROBABILITY	0.0168
ALIENATION	0.84564

CORRELATION BETWEEN THE VARIABLES AND THE CANONICAL VARIATES WITHIN THE HYPOTHESIS SET

VARIABLE	CANONICAL VARIATE 1	2	3	4	5	6	7	8	9
1	0.2994	0.2350	-0.0531	-0.0[illegible]22	0.0204	0.0044	0.1071	0.0317	0.0247
2	-0.0734	-0.0442	0.1920	0.0403	0.0039	-0.0063	-0.0216	-0.0059	-0.0117
3	0.0555	0.3789	0.4393	0.0217	0.0287	0.0280	0.0415	0.0237	-0.0268
4	-0.4103	-0.1007	0.5491	0.1838	-0.0006	0.0062	-0.1440	-0.0[illegible]	-0.0699
5	0.2168	0.0860	0.2650	-0.0006	0.0300	-0.0190	0.0688	0.0242	0.0065
6	-0.0869	0.6326	0.0511	0.0062	-0.0190	0.1212	-0.0655	-0.0064	-0.0651
7	0.5697	-0.1697	0.3994	-0.1440	0.0688	-0.0555	0.2715	0.[illegible]74	0.07[illegible]
8	-0.0816	0.3918	-0.1258	-0.0305	0.0242	-0.0064	0.0774	0.0254	0.0113
9	0.6755	-0.5995	-0.0501	-0.0690	0.0065	-0.0651	0.0784	0.0113	0.0596

STANDARDIZED SCORE WEIGHTS FOR REGRESSING VARIABLES ONTO CANONICAL VARIATES WITHIN THE HYPOTHESIS SET

VARIABLE	CANONICAL VARIATE 1	2	3
1	0.63780	0.63612	-0.[illegible]
2	-0.13132	-0.18531	0.1[illegible]
3	-0.20367	0.13616	0.57741
4	-0.30521	-0.21744	0.77378
5	0.10801	0.16567	0.98980
6	0.29537	0.43913	-0.50329
7	0.77108	0.24313	0.45750
8	0.19214	0.42204	0.36170
9	0.39097	-0.59612	-0.22488

RAW SCORE WEIGHTS FOR REGRESSING VARIABLES ONTO CANONICAL VARIATES WITHIN THE HYPOTHESIS SET

VARIABLE	CANONICAL VARIATE 1	2	3
1	0.01428	0.01429	-0.00221
2	-0.00392	-0.00547	0.00359
3	-0.00747	0.00500	0.01361
4	-1.15392	-0.82208	2.92515
5	0.42943	0.65066	3.93512
6	1.50091	2.23143	-2.55745
7	0.28959	0.09141	0.17162
8	0.06226	0.15679	0.12572
9	0.19251	-0.29555	-0.11175

WEIGHTS FOR REGRESSING RAW SCORE HYPOTHESIS VARIABLES WITH HYPOTHESIS CANONICAL VARIATES
WITH UNIT SUMS-OF-SQUARES

VARIABLE	CANONICAL VARIATE 1	2	3
1	0.90302	0.90334	-0.13946
2	-0.32940	-0.45979	0.33567
3	-0.53229	0.35587	0.98377
4	-0.30521	-0.21744	0.77370
5	0.10801	0.16367	0.98980
6	0.29537	0.43913	-0.50329
7	0.77108	0.24313	0.45750
8	0.19214	0.42204	0.38170
9	0.39097	-0.59612	-0.22488

ANALYSIS OF CORRELATION

HYPOTHESIS TESTED V2=V1

CORRELATIONS AMONG THE HYPOTHESIS VARIABLES

VARIABLE	ERROR 4	ERROR 5	ERROR 6
ERROR 4	1.00000	0.07416	-0.11655
ERROR 5	0.07416	1.00000	-0.44022
ERROR 6	-0.11655	-0.44022	1.00000

STANDARD DEVIATIONS OF HYPOTHESIS VARIABLES

VARIABLE	ERROR 4	ERROR 5	ERROR 6
	127.22711	0.89856	9.40055

CROSS CORRELATIONS BETWEEN HYPOTHESIS AND ERROR VARIABLES

VARIABLE	ERROR 4	ERROR 5	ERROR 6
ERROR 1	0.08668	0.32015	-0.10542
ERROR 2	-0.16004	-0.15982	-0.07656
ERROR 3	0.02616	0.09527	-0.32799

RAW SCORE WEIGHTS FOR REGRESSING HYPOTHESIS VARIABLES ONTO ERROR VARIABLES

ERROR	HYPOTHESIS ERROR 4	ERROR 5	ERROR 6
ERROR 1	0.08400	59.40628	0.85901
ERROR 2	-0.02609	-5.24144	-0.42469
ERROR 3	-0.00205	-1.62598	-0.91256

STANDARD SCORE WEIGHTS FOR REGRESSING HYPOTHESIS VARIABLES ONTO ERROR VARIABLES

ERROR	HYPOTHESIS ERROR 4	ERROR 5	ERROR 6
ERROR 1	0.06759	0.33760	0.05107
ERROR 2	-0.16584	-0.23534	-0.19949
ERROR 3	-0.01083	-0.06062	-0.35593

STANDARD ERROR ABOUT THE REGRESSION LINE FOR ERROR VARIABLES

VARIABLE	ERROR 1	ERROR 2	ERROR 3
	158.34528	20.36580	24.10719

CORRELATIONS AMONG ERROR VARIABLES

VARIABLE	ERROR 1	ERROR 2	ERROR 3
ERROR 1	1.00000	-0.14650	-0.24907
ERROR 2	-0.14650	1.00000	0.47822
ERROR 3	-0.24906	0.47823	1.00000

STANDARD DEVIATIONS OF THE ERROR VARIABLES

VARIABLE	ERROR 1	ERROR 2	ERROR 3
	158.11751	20.01227	24.10139

REGRESSION SUMS-OF-SQUARES WHEN TOTAL SUMS-OF-SQUARES OF ERROR VARIABLES ARE UNITIES

VARIABLE	ERROR 1	ERROR 2	ERROR 3
ERROR 1	0.10855	-0.06868	0.01718
ERROR 2	-0.06868	0.07943	0.03867
ERROR 3	0.01718	0.03867	0.11068

STATISTICAL SUMMARY

F-RATIO FOR WILKS LAMBDA CRITERION	0.945[illegible]
NUMERATOR DEGREES OF FREEDOM	9.
DENOMINATOR DEGREES OF FREEDOM	5[illegible].7
PROBABILITY	0.4944

DIMENSION REDUCTION STATISTICS FOR LARGE F. BERF, N= 24.

CANONICAL VARIATE	1	2	3
CANONICAL CORRELATIONS	0.4692	0.3071	0.1017
INDIVIDUAL LAMBDAS	0.2[illegible]2	0.104	0.010
CHI-SQUARE FOR ROOTS K THROUGH LAST	7.9606	2.4554	0.2312
DEGREES OF FREEDOM	9.	4.	1.
PROBABILITY	0.5367	0.6557	0.6308

UNIVARIATE F STATISTICS FOR ERROR VARIABLES, 3 AND 24 DEGREES OF FREEDOM

VARIABLE	ERROR 1	ERROR 2	ERROR 3
MULTIPLE CORRELATION	0.3295	0.2818	0.3327
F-RATIO	0.97413	0.69025	0.99567
PROBABILITY	0.4213	0.5669	0.4117
ALIENATION	0.94417	0.95946	0.94304

CORRELATIONS BETWEEN THE VARIABLES AND THE CANONICAL VARIATES WITHIN THE ERROR SET

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 1	0.6309	-0.4181	0.6555
ERROR 2	-0.3031	0.7702	0.5612
ERROR 3	0.3889	0.9039	-0.1783

STANDARD SCORE WEIGHTS FOR REGRESSING VARIABLES ONTO CANONICAL VARIATES WITHIN THE ERROR SET

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 1	0.75819	-0.19317	0.67458
ERROR 2	-0.60714	0.43131	0.86269
ERROR 3	0.86811	0.64948	-0.42257

RAW SCORE WEIGHTS FOR REGRESSING VARIABLES ONTO CANONICAL VARIATES WITHIN THE ERROR SET

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 1	0.00480	-0.00122	0.00427
ERROR 2	-0.03034	0.02155	0.04308
ERROR 3	0.03602	0.02695	-0.01753

WEIGHTS FOR REGRESSING RAW SCORE ERROR VARIABLES ONTO ERROR CANONICAL VARIATES WITH UNIT SUMS-OF-SQUARES

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 1	0.00700	-0.00220	0.01338
ERROR 2	-0.04429	0.03889	0.13510
ERROR 3	0.05258	0.04863	-0.05499

REGRESSION SUMS-OF-SQUARES WHEN TOTAL SUMS-OF-SQUARES OF HYPOTHESIS VARIABLES ARE UNITIES

VARIABLE	ERROR 4	ERROR 5	ERROR 6
ERROR 4	0.04738	0.07971	-0.04366
ERROR 5	0.07971	0.18465	-0.11881
ERROR 6	-0.04366	-0.11881	0.15217

UNIVARIATE F STATISTICS FOR HYPOTHESIS VARIABLES, 3 AND 24 DEGREES OF FREEDOM

VARIABLE	ERROR 4	ERROR 5	ERROR 6
MULTIPLE CORRELATION	0.2177	0.4297	0.3901
F-RATIO	0.39788	1.81179	1.43584
PROBABILITY	0.7557	0.1719	0.2570
ALIENATION	0.97602	0.90297	0.92078

CORRELATION BETWEEN THE VARIABLES AND THE CANONICAL VARIATES WITHIN THE HYPOTHESIS SET

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 4	0.3955	0.2240	-0.8907
ERROR 5	0.9003	0.2244	0.3750
ERROR 6	-0.6781	0.7348	0.0145

STANDARD SCORE WEIGHTS FOR REGRESSING VARIABLES ONTO CANONICAL VARIATES WITHIN THE HYPOTHESIS SET

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 4	0.30377	0.29832	-0.91278
ERROR 5	0.73786	0.67107	0.49640
ERROR 6	-0.31786	1.06503	0.12664

RAW SCORE WEIGHTS FOR REGRESSING VARIABLES ONTO CANONICAL VARIATES WITHIN THE HYPOTHESIS SET

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 4	0.00239	0.00234	-0.00717
ERROR 5	0.82116	0.74683	0.55243
ERROR 6	-0.03381	0.11329	0.01347

WEIGHTS FOR REGRESSING RAW SCORE HYPOTHESIS VARIABLES ONTO HYPOTHESIS CANONICAL VARIATES
WITH UNIT SUMS-OF-SQUARES

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 4	0.64739	0.63577	-1.94527
ERROR 5	2.40262	2.18513	1.61655
ERROR 6	-3.12645	10.47571	1.24567

VARIMAX ROTATION

DECISION CRITERION FOR CANONICAL VARIATES: P LESS THAN 0.750
TAXONOMY OF VARIABLES
DIRECT ROTATION

INTERMEDIATE OUTPUT

CYCLE	CRITERION	DIFFERENCE	STAND. CRIT.	STAND. DIF.	NUM. ROT.	NUM. 0 ROT.
1	1.68060	1.68060	0.84030	0.84030	3	0
2	1.70725	0.02665	0.85362	0.01332	6	0
3	1.70726	0.00001	0.85363	0.00000	7	2
4	1.70726	0.0	0.85363	0.0	7	5

CORRELATIONS BETWEEN THE ERROR VARIABLES AND THE ROTATED CANONICAL VARIATES

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 1	0.9916	-0.0594	-0.1149
ERROR 2	-0.0618	0.9685	0.2473
ERROR 3	-0.1250	0.2464	0.9611

TRANSFORMATION MATRIX

CAN. VAR.	TRANS. VECTOR 1	2	3
	0.6808	-0.4192	0.6007
	-0.2994	0.5892	0.7505
	0.6685	0.6907	-0.2756

SUM-OF-SQUARES MATRIX OF ROTATED HYPOTHESIS CANONICAL VARIATES

	CANONICAL VARIATE 1	2	3
1	0.1151	-0.0747	0.0669
2	-0.0747	0.0764	-0.0157
3	0.0669	-0.0157	0.1333

WEIGHTS FOR REGRESSING ERROR VARIABLES ONTO ROTATED ERROR CANONICAL VARIATES

VARIABLE	CANONICAL VARIATE 1	2	3
ERROR 1	0.0065	0.0002	0.0008
ERROR 2	0.0017	0.0552	-0.0139
ERROR 3	0.0047	-0.0113	0.0467

TEST WITH NO ERROR NOT DONE
N=0

TEST WITH NO ERROR NOT DONE
N=0

VARAN: LINEAR MODEL VARIANCE ANALYSIS

FIRST EDITION: JUNE 1, 1972

EDUCATIONAL TESTING SERVICE

PRINCETON, N.J. 08540

END OF PROBLEMS

Selected Bibliography

1. General linear model

   a. Scheffe, H. The Analysis of Variance. New York: Wiley, 1959.

   b. Bock, R. D. Programming univariate and multivariate analysis of variance. Technometrics 5: 95-117, 1963.

   c. Wilks, S. S. Certain generalizations in the analysis of variance. Biometrika 24: 471-494, 1932.

   d. Wilks, S. S. The analysis of variance and covariance in nonorthogonal data. Metron 13: 141-154, 1938.

2. Canonical correlation

   a. Hotelling, H. Relations between two sets of variates. Biometrika 28: 321-377, 1936.

   b. Meredith, W. Canonical correlations with fallible data. Psychometrika 29: 55-66, 1964.

3. Discriminant analysis and multivariate analysis of variance

   a. Hall, C. E. Three Papers in Multivariate Analysis: I, Interpretation of multivariate analysis with an example in multivariate analysis of variance. Pittsburgh, PA: American Institutes for Research, 1967.

   b. Morrison, D. F. Multivariate Statistical Methods. New York: McGraw-Hill, 1967.

   c. Rao, C. R. Advanced Statistical Methods in Biometric Research. New York: Wiley, 1953.

   d. Rao, C. R. Linear Statistical Inference and Its Applications. New York: Wiley, 1965.

4. Miscellany

a. Analysis of covariance

Hall, C. E. Analysis of covariance for nonrandomized data. Research Bulletin 71-27. Princeton, N. J.: Educational Testing Service, 1971.

b. Significance test card notation

Hall, C. E. A symbolic logic for representing linear models. Research Bulletin 71-60. Princeton, N. J.: Educational Testing Service, 1971. Also presented at the American Statistical Association Meetings, Montreal, Canada, September 1972.

c. Rotation

(1) Hall, C. E. Rotation of canonical variates in multivariate analysis of variance. Journal of Experimental Education 38 (No.2): 31-38, 1969.

(2) Hall, C. E. A factor rotation scheme giving orthogonal bounds. Journal of Experimental Education 40 (No. 2): 49-50, 1971. (Also Research Bulletin 71-24. Princeton, N. J.: Educational Testing Service, 1971.)

(3) Cliff, N. Orthogonal rotation to congruence. Psychometrika 31(No. 1), 33-42.

d. Battery reduction

Hall, C. E. Generalizing the Wherry-Doolittle battery reduction procedure to canonical correlation and MANOVA. Journal of Experimental Education 39 (No. 4): 47-51, 1971.

e. Extension

Hall, C. E. Extension of discriminant analysis and multivariate analyses of variance. In Proceedings of the 76th Annual Convention of the American Psychological Association. Washington, D.C.: American Psychological Association, 1969.

f. Factor analysis

(1) Browne, M. Gauss-Seidel computing procedures for a family of factor analytic solutions. Research Bulletin 68-61. Princeton, N. J.: Educational Testing Service, 1968.

(2) Harris, C. W. Some Rao-Guttman relationships. Psychometrika 27: 247-263, 1962.

(3) Harman, H. H. Modern Factor Analysis (2nd ed.) Chicago: University of Chicago Press, 1967.

g. Dimension reduction

Hall, C. E. Dimension reduction analysis. Research Bulletin 72-55. Princeton, N. J.: Educational Testing Service, 1972.

h. Variance reduction

Hall, C. E. Variance reduction analysis. Research Bulletin 73-29. Princeton, N. J.: Educational Testing Service, 1973.

i. Serial correlation

Hall, C. E. Studies of single samples and whole populations. Research Bulletin 73-44. Princeton, N. J.: Educational Testing Service, 1973.

j. Potthoff and Roy models

Potthoff, R. F. and Roy, S. N. A generalized MANOVA model useful especially for growth curve problems. Biometrika 51: 315-328, 1964.

Grizzle, J. E. and Allen, D. N. Analysis of growth and response curves. Biometrics, 357-381, 1969.

Khatri, C. G. A notion a MANOVA model applied to problems in growth curves. Tokyo: Annals of the Institute of Statistical Mathematics, 18: 75-86, 1966.